AD-A258 682

USACERL Technical Manuscript FF-93/03
November 1992

US Army Corps
of Engineers
Construction Engineering
Research Laboratory

Least Square Methods for Solving Systems of Inequalities With Application to an Assignment Problem

by
Janet H. Spoonamore

This research addresses algorithmic approaches for solving two different, but related, types of optimization problems. Firstly, the research considers the solution of a specific type of assignment problem using continuous methods. Secondly, the research addresses solving systems of inequalities (and equalities) in a least square sense. The specific assignment problem has piece-wise linear additive separable server cost functions, which are continuous everywhere except at zero, the point of discontinuity for the {0,1} assignment condition. Continuous relaxation of the {0,1} constraints yields a linear programming problem. Solving the dual of the linear programming problem yields the complementarity conditions for a primal solution, a system of linear inequalities and equalities. Adding equations to this system to enforce a {0,1} solution in the relaxed solution set yields an augmented system, not necessarily linear. Methods to solve this system, a system of linear inequalities and non-linear equations, in a least square sense are developed. The specific assignment problem is a variation of problems which are amenable to strong continuous relaxation, in that the solution set of the relaxed problem has been shown, experimentally, to often contain a {0,1} solution. However, if there are a large number of variables, efficient continuous (non-combinatoric) methods are needed to locate {0,1} solutions, if such exist. This work addresses methods to find {0,1} solutions using a least square formulation for solving systems of inequalities.

There is a large body of work dealing with nondifferentiable optimization, but the kind of nondifferentiability posed by inequalities is of a special type. By considering a least square formulation of the problem:

(a) meaningful solutions can be found even in the infeasible case which naturally arises in applications such as sequential quadratic programming;
(b) the problem becomes one (but not twice) differentiable;
(c) the generalized second order differential set has favorable properties that allow generalization of classical second order non-linear least square algorithms;
(d) the fundamental computational subproblem, solving a system of linear equalities, is efficiently solvable;
(e) the problem becomes numerically stable (in the sense of Robinson).

Common algorithmic approaches to solve nonlinear least square problems are adapted to solve systems of inequalities. Local and global convergence results are developed, using properties of the Clarke generalized subdifferential and Jacobian. Rates of convergence are analyzed. Applications of the algorithms for solving the piece-wise linear assignment subproblem are developed and analyzed. Application of the algorithms for solving linear programming problems, and linear and convex complementarity problems are described.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave Blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	November 1992	Final

4. TITLE AND SUBTITLE
Least Square Methods for Solving Systems of Inequalities With Application to an Assignment Problem

5. FUNDING NUMBERS
n/a

6. AUTHOR(S)
Janet H. Spoonamore

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
U.S. Army Construction Engineering Research Laboratories (USACERL)
PO Box 9005
Champaign, IL 61826-9005

8. PERFORMING ORGANIZATION REPORT NUMBER
TM FF-93/03

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
n/a

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES
Copies are available from the National Technical Information Service, 5285 Port Royal Road, Springfield, VA 22161.

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

This research addresses algorithmic approaches for solving two related types of optimization problems: (1) solving a specific type of assignment problem using continuous methods and (2) solving systems of inequalities (and equalities) in a least square sense. The specific assignment problem has piece-wise linear additive separable server cost functions, which are continuous everywhere except at zero, the point of discontinuity for the {0,1} assignment condition. Continuous relaxation of the {0,1} constraints yields a linear programming problem. Solving the dual of the linear programming problem yields the complementarity conditions for a primal solution, a system of linear inequalities and equalities. Adding equations to this system to enforce a {0,1} solution in the relaxed solution set yields an augmented system, not necessarily linear. Methods to solve this system, a system of linear inequalities and non-linear equations, in a least square sense are developed.

Common algorithmic approaches to solve nonlinear least square problems are adapted to solve systems of inequalities. Local and global convergence results are developed. Rates of convergence are analyzed. Applications of the algorithms for solving the piece-wise linear assignment subproblem are developed and analyzed. Application of the algorithms for solving linear programming problems, and linear and convex complementarity problems are described.

14. SUBJECT TERMS
assignment problem, Clark subdifferential, least square problem, global convergence, stability, almost everywhere differentiable, local convergence

15. NUMBER OF PAGES
82

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	SAR

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std 239-18
298-102

FOREWORD

This manuscript was submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Mathematics in the Graduate College of the University of Illinois at Urbana-Champaign. The University of Illinois advisor for this thesis was Dr. Randall Bramley.

Janet Spoonamore is a principal investigator for the Facility Management Division (FF) of the Infrastructure Laboratory (FL), U.S. Army Construction Engineering Research Laboratories (USACERL). Research for this manuscript was conducted during a fully funded academic leave from USACERL. Alan Moore is Acting Chief, Facility Management Division (CECER-FF). Dr. Michael J. O'Connor is Chief, CECER-FL.

COL Daniel Waldo, Jr., is Commander and Director of USACERL, and Dr. L.R. Shaffer is Technical Director.

Table of Contents

Notation

1) $\Re^n$ denotes the real n-dimensional space. Let 1^n denote the vector of all ones in $\Re^n$.

2) $\Re^n_+$ denotes the positive orthant in $\Re^n$.

3) Let S be a subset of $\Re^n$.

   - convex hull S denotes the set of all finite convex combinations of elements of S.

   - $\{d \mid x \in S$ implies $x + \lambda d \in S$ for all $\lambda \geq 0\}$ is the set of directions of recession of S.

   - Let $\delta_S(x)$ be the indicator function of the set, S. $\delta_S(x) := 0$, if $x \in S$; otherwise, $\delta_S(x) := +\infty$.

4) Let $\langle \bullet, \bullet \rangle$ denote the inner product on $\Re^n$. Let $S^\perp := \{y \mid \langle x, y \rangle = 0$, for all $x \in S\}$ denote the orthogonal complement of the set S. Given $a \in \Re^n$, a scalar α, let $H := \{x \mid \langle a, x \rangle = \alpha\}$ denote the hyperplane associated with a and α.

5) Let $K \subset \Re^n$ be a convex cone, then denote $K^* := \{y \mid \langle x, y \rangle \leq 0$ for all $x \in K\}$. K^* is called the polar of the cone K.

6) Let $C^1[\Re^n, \Re^m]$ be the set of all continuously once differential functions from $\Re^n$ to $\Re^m$. Similarly, $C^2[\Re^n, \Re^m]$ is the set of all continuously twice differential functions from $\Re^n$ to $\Re^m$.

7) Let $x \in \Re^n$. Let $S \subset \Re^n$. All vectors are written as column vectors.

   - $\|x\|$ denotes the Euclidean norm of x.

   - dist $(x, S) := \inf\{\|x - y\| \mid y \in S\}$.

   - $N(x, r) := \{y \mid \|x - y\| < r\}$ denotes an open neighborhood of radius r about x.

   - x_+ denotes the Euclidean projection of x onto $\Re^n_+$.

   - $x_- := x + x_+$.

8) For a linear map G,

   - denote $\|G\|$ as the induced operator norm,

   - denote $\mathcal{N}$(G) as the null space of G.

9) A function f is Lipschitz of order p of rank $\gamma \geq 0$ on a set S, if all $x, y \in S$ satisfy: $\|f(x) - f(y)\| \leq \gamma\|x - y\|^p$.

10) Let A be an m by n matrix and let $I \subset \{1, 2, \ldots, m\}$.

   - Denote $|I|$ to mean the cardinality of the set I.

   - A_I denotes the p by n matrix, where $p = |I|$, formed from the p rows of A indexed by I.

   - $A^\dagger$ denotes the pseudo-inverse of the matrix A, the generalized inverse giving a least norm, least square solution.

   -Typically, denote index i, as the row of a matrix.

   -Typically, denote index j, as the column of a matrix.

11) For a matrix A, an n by n matrix, denote $\|A\|_F$, the Frobenius norm of A,

$$\|A\|_F := \left(\sum_{j=1}^{n} \sum_{i=1}^{n} |a_{i,j}|^2 \right)^{\frac{1}{2}}.$$

12) Let S be a subset of $\Re^n$. Usually f is a single-valued functional mapping of S into $\Re$, and F is a single-valued mapping from S into $\Re^m$. A property of f is satisfied a.e. on S, if this property is satisfied everywhere on S except on a set of measure 0.

13) The Clarke subdifferential set of f at x is denoted $\partial f(x)$. The gradient of f at x is denoted $\nabla f(x)$.

14) The Clarke generalized Jacobian set of F at x is donoted $\partial F(x)$. The Jacobian of $F(x)$ is denoted $\mathcal{J}F(x)$. Similarly, the Clarke generalized Jacobian set of $\nabla f(x)$ (sometimes referred to as a generalized Hessian) is denoted $\partial \nabla f(x)$. The Hessian of f at x is denoted $\nabla^2 f(x)$.

15) In algorithms, iterates are typically denoted with the index, k. Greek lower case letters, such as $\alpha, \beta, \gamma, \lambda$, are usually positive scalars. Direction vectors are usually denoted d, and the algorithm scaled step is usually s. Distinquished points, such as solution or accumulation points are usually denoted by a star convention, such as x^*.

16) The notation, "little o" means $\alpha = o(\delta) \iff \frac{\alpha}{\delta} \to 0$ as $\delta \to 0$.

17) $\triangle f_k$ and $\triangle M_k$ are symbols used in development of trust region algorithms, where at each iterate k, $\triangle f_k := f(x_k) - f(x_k + s_k)$ and $\triangle M_k := M_k(0) - M_k(s_k)$.

1 Introduction

1.1 Types of Optimization Problems

Mathematical optimization methods include two distinct types, the combinatoric and the continuous methods. Unfortunately, practicing operations researchers often face problems which are neither purely continuous nor combinatoric. For these problems, one cannot directly use the strong algorithms for continuous optimization, nor the discrete methods. One such hybrid problem is the assignment problem, a useful formulation for cost minimization problems. Assignment problems involve assigning jobs to servers, where one wishes to match servers and jobs based on minimizing total costs. For a linear assignment, each server cost function is linear. Realistic assignment problems cannot be adequately modeled by formulation as a linear assignment problem, a computationally solvable problem. Practicing operations researchers look for methods allowing representation of non-linear properties, when the server cost functions are highly non-linear. For example, piece-wise linear approximations offer greater accuracy. Another typical modeling requirement may be satisfied by approximation using a cost function having a flat minimum charge and linear rates beyond the minimum charge. The added accuracy gained by allowing piece-wise linear server cost functions, even just two pieces, is an attractive capability. The assignment problem having additively separable piece-wise linear server cost functions can be solved by a strong relaxation into a linear program and then using Gomory cutting plane methods for solution of $\{0, 1\}$ assignment variables. However, these combinatoric methods are unsatisfactory for large scale problems. Recent work addressing improved large scale methods include the work of Conn and Cornuéjols[1990], where the problem is simplified by solving for the unconstrained dual using a projected gradient to derive search directions. Their work does not address specific methods for finding $\{0, 1\}$ primal solutions generated by a dual solution, which is formidable in the case of large scale problems. We develop a method, based on an algorithm for finding a least square solution of a system of linear inequalities and non-linear equations, along with using common Gomory cutting planes to solve for the $\{0, 1\}$ solutions.

Recent work addressing least square methods for solving systems of inequalities centers on the research of Han[1980], Burke[1983], and Han and Burke [1986]. Han[1980] developed an algorithm for solving systems of linear inequalities using a search direction derived as the projection of the negative of the residuals onto the subspace spanned by the coefficients of the unmet and active constraints. Further, he showed the existence of a neighborhood of any solution, where the active and unmet constraints of the current iterate are a subset of the active and unmet constraints of the next iterate. This neighborhood is now called an identification neighborhood. This property assured finite termination of his algorithm.

In Burke[1983] and Burke and Han[1986], a function property called K-regular is defined, which guarantees a well-defined search direction at each iterate. While this class of functions may seem quite general, Burke and Han[1986] cite "seemingly well-behaved situations" which are excluded from their K-regular class of functions. Using this property, they extended the least square algorithm for solving finite dimensional systems of inequalities and equalities of K-regular functions using the same type of projection as for the linear case. Burke[1983] showed sufficiency conditions for global convergence of this "Gauss-Newton" type search direction, as well as the steepest descent direction, using Armijo line search conditions to solve systems of inequalities of uniformly K-regular C^1 functions in a least square sense. For the non-linear case, they used the Mangasarian Fromowitz constraint qualification condition, defined in section 6, to assure local q-quadratic convergence in the special identification neighborhood of a stationary point having zero residual.

This dissertation addresses solving finite systems of inequalities of C^2 convex functions and equalities of C^2 functions, in a least square sense. The advantage of first examining the class of convex functions is that the strong theoretical properties are revealed, suggesting the needed properties to generalize for the non-convex case. Further, for the assignment application subproblem addressed here, the inequalities satisfy this C^2 convex condition. This dissertation extends, to the inequality case, the traditional methods used to solve the equality case, including the Gauss-Newton, Levenberg-Marquardt, Newton's methods, and quasi-Newton updates. We derive new adaptations of these methods to handle solving inequalities in a least square sense. We show the sufficient local convergence properties to assure various convergence rates, such as q-quadratic, or q-superlinear, or q-linear convergence. Further, we analyze and adapt the other major global convergence method, the trust region method, to solve systems of inequalities in a least square sense. We show sufficient properties for global convergence of our modified trust region algorithm to handle solving systems of inequalities in a least square sense.

Another important result of this dissertation is that we have unified and greatly simplified the treatment of the essential properties which govern global and local convergence when solving systems of inequalities in a least square sense. For local convergence, the key property of the least square objective function which unifies the underlying theory, regardless of the algorithm, is the special approximating property that the generalized Clarke differential of the least square objective function possesses in the special identification neighborhood of a designated point. This special approximating property assures that the second order Taylor series expansion (based on generalized differential constructs) about each point x, sufficiently close to a given point x_*, models the objective function at x_* perfectly in the case of linear inequalities and with second order accuracy in the nonlinear convex case. In contrast, a "true" analytic property would also assure that the expansion about x_* models each point x. The key global property is that the gradient of the least square objective function for inequalities is uniformly Lipschitz of some rank γ on the lower level set of the first iterate. The model trust region method, as with the Armijo line search, is shown to converge without any added conditions, because of the uniform Lipschitzian property of the gradient of the objective function. With our results, we demonstrate that one can use existing algorithms, with minor adaptations, to solve for least square solutions of systems of convex inequalities and non-linear equations.

1.2 Overview of Thesis

Section 1.3 in the Introduction gives background material on generalized differentiability, as in Clarke[1989], as well as other results needed for the analysis of algorithm convergence properties. These properties are the foundation of the dissertation development. Section 1.4 gives a description of a somewhat remarkable, yet simple, algorithm developed by S.P. Han to solve linear systems of inequalities in a least square sense. The method to find search directions in Han's algorithm uses a singular value decomposition (SVD). Using the result of Bramley[1992], that the search directions generated by QR factorization with column pivoting are bounded by the minimum norm solution provided by SVD, it follows that the method also converges using the easier-to-compute search direction using a QR factorization with column pivoting. Since the QR factorization with column pivoting is a finite computational process, and SVD is a convergent iterative process, this change allows the method to be used as a practical algorithm. The ideas in this algorithm for linear inequalities are a motivation to consider extension to other common minimization algorithms.

Section 1.5 in the Introduction describes background on algorithms for solving systems of equations in a least square sense.

Section 2 develops the assignment problem, starting with the original primal problem in $\{0, 1\}$ variables, strong relaxation to a linear programming primal, conversion to the relaxed dual, and characterization of relaxed primal solutions from a dual solution. An example problem is given in section 2.4, and section 2.5 discusses methods to solve for $\{0, 1\}$ solutions within the solution set generated by a relaxed dual solution.

The treatment of least square solution of inequalities and equalities is done in sections 3 and 5. In section 3.1, we develop the least square formulation for solving systems of linear inequalities and nonlinear equations. In section 3.2, we state the definition of stability of a system of inequalities, based on Robinson[1976]. We show that the least square formulation for solving a system of inequalities is numerically stable to small perturbations or inaccuracies in computation, using this definition of stability. In this section, we give an example of a system of inequalities which is not stable, according to Robinson's definition of stability; but the formulation of this system as a least square problem yields a system of inequalities which is stable.

In section 3.3, the identification neighborhood of a point is defined. The special approximating property of the generalized differential constructs of the least square objective function is defined and verified for the case of linear inequalities. In this case, the Taylor series based on second order generalized differential construct models the least square objective function perfectly in an identification neighborhood.

In section 4.1, we develop the Gauss-Newton and Levenberg-Marquardt like algorithms and show local convergence properties. We verify global convergence properties for the Armijo line search and develop and analyze global trust region methods using a generalized Hessian model function in section 4.2. In section 4.3, we analyze the least square objective function properties, which assure that the iterates, for which function decreases, stay bounded. This assures the existence of an accumulation point among the iterates. This function property is related to coercive behavior, which is defined in this section.

In section 5, we develop the case of systems of inequalities of convex C^2 functions and nonlinear equations. In this section, we show the special approximating property of the Taylor series based on generalized differential second order constructs of the least square objective function gives second order approximation in an identification neighborhood of a point. From this, in section 6, we are able to develop and prove local convergence for the Gauss-Newton like algorithm, Levenberg-Marquardt like algorithm, the Newton-like algorithm and the special quasi-Newton symmetric secant update algorithm for non-linear least square problems. We discuss global convergence results, noting the similarity to linear case.

In section 7, we develop and verify applications, specifically the solution of the assignment problem of section 2. We develop and analyze a method to solve this problem, based on Han's algorithm, and on the algorithm for solving systems of non-linear equations and linear inequalities in a least square sense. In section 7.2, we show the application of using the least squares method to solve linear programs, specifically, the relaxed linear program of the assignment problem. Thus, we show how one could use this method to solve both subproblems of the assignment problem. In section 8, examples of the Gauss-Newton like algorithm for solving the sample piece-wise linear assignment problem in section 2 are shown.

In section 9, conclusions and recommendations for further study are given.

1.3 Background on Generalized Differentials

We use properties of the Clarke[1980] generalized Jacobian and subgradient in this research. Given a function, F, where $F : \Re^n \mapsto \Re^m$, define F to be *Lipschitz of rank* γ *of order 1* on some neighborhood of x, if for all y in this neighborhood,

$$\|F(x) - F(y)\| \leq \gamma \|x - y\|.$$

The Clarke subgradient set is defined for finite dimensional functions, which is convenient for calculations. Let $f : \Re^n \mapsto \Re$ be Lipschitz of order 1 of rank γ in a neighborhood of x . Let S be a set of Lebesque measure 0 in $\Re^n$, and let Ω be the set of points where f fails to be differentiable. By Rademacher's Theorem, Ω has measure 0. Then the *Clarke generalized subgradient* of f at x is the set $\partial f(x)$, where

$$\partial f(x) = \text{ convex hull } \{\lim \nabla f(x_l) \,|\, x_l \rightarrow x, x_l \notin S, x_l \notin \Omega\}.$$

If f is a finite convex function, then the Clarke generalized subgradient set is the same as the convex subgradient set defined as:

$$\partial f(x) := \{x^* | \, f(y) - f(x) \leq \langle x^*, y - x\rangle, \text{ for all } y \in \Re^n\}.$$

Now, consider $F : \Re^n \mapsto \Re^m$, and assume that F is Lipschitz of rank γ of order 1 on some neighborhood of x. Let Ω be the set of points where F fails to be differentiable. Denote the Jacobian of F at x_l as $\mathcal{J}F(x_l)$, for $x_l \notin \Omega$, where $l \in I$, some index set. Then the *Clarke generalized Jacobian* of F at x is the set $\partial F(x)$, where

$$\partial F(x) := \text{ convex hull } \{\lim \mathcal{J}F(x_l) \,|\, x_l \rightarrow x, x_l \notin \Omega, l \in I\}.$$

$\partial F(x)$ is a set of all $m \times n$ matrices which are convex combinations of a finite number of limits of a sequence of Jacobians, $\mathcal{J}F(x_l)$, as x_l approaches x.

Some important properties of the generalized subgradient and Jacobian:

1. $\partial f(x)$ is a nonempty convex compact subset of $\Re^n$.
2. $\partial F(x)$ is a nonempty convex compact subset of $\Re^{mn}$.
3. If f is differentiable at x, then $\partial f(x) = \nabla f(x)$.
4. If F is differentiable at x, then $\partial f(x) = \mathcal{J}F(x)$.
5. ∂F is an upper semicontinuous mapping at x. That is, for any $\delta > 0$, there is a $\epsilon > 0$, such that for $y \in N(x, \epsilon)$, a neighborhood of x, then:

$$\partial F(y) \subset \partial F(x) + \delta B_{mn},$$

   where B_{mn} denotes the unit ball in $\Re^{mn}$.

6. *The Vector Mean Value Theorem for Generalized Jacobian*: Let F be Lipschitz on an open convex set U in $\Re^n$, and let $u, v \in U$. Then one has:

$$F(v) - F(u) \in \text{ convex hull } \partial F([u, v])(v - u),$$

   where $[u, v]$ represents the convex hull of $\{u, v\}$.

7. *The Jacobian Chain Rule*: Let $F : \Re^n \mapsto \Re^m$ be Lipschitz in a neighborhood of x and $G : \Re^m \mapsto \Re^k$ be Lipschitz near $F(x)$. Then $G \circ F$ is Lipschitz in some neighborhood of x and for any $p \in \Re^n$, one has

$$\partial(G \circ F(x))p \subset \text{ convex hull } \{\partial G(F(x))\partial F(x)p\},$$

   and, if G is strictly differentiable, one has:

$$\partial(G \circ F(x))p = \nabla G(F(x))\partial F(x)p.$$

Note that $\partial F(x)$ is a set of $m \times n$ matrices, since $F(x)$ is a map from $\Re^n \mapsto \Re^m$, and $\partial G(F(x))$ denotes the generalized Jacobian of G with respect to $F(x)$, so that $\partial G(F(x))$ is a set of $k \times m$ matrices, and, thus, $\partial(C \circ F(x))$ is a set of $k \times n$ matrices.

Some notational details are given here. In our discussions we compare the properties of Clarke generalized differentiability versus the common notion of differentiability. For a functional f, first and second order generalized differentiability properties are discussed. For a function F, first order generalized differentiability is discussed. For ordinary differentiable constructs, use the symbol ∇ for the gradient, the symbol ∇^2 for the Hessian, and use the caligraphic $\mathcal{J}$ symbol for the Jacobian. We reserve the use of the notation ∂f and ∂F to mean the Clarke generalized subgradient set and generalized Jacobian set, respectively, throughout the dissertation unless otherwise noted. Suppose f_1 is twice differentiable. Suppose f_2 is once-differentiable and ∇f_2 is Lipschitz continuous of rank γ on some neighborhood of x which assures that f_2 is second order generalized differentiable at x. Suppose F_1 is differentiable at x, and F_2 is Lipschitz continuous of rank γ on some neighborhood of x, thus assuring F_2 is first order generalized differentiable at x. The notation for symbols used to denote differentiable and generalized differentiable constructs is summarized in Table 1, Notation for Differential Constructs, below:

Order	Differential Construct	Generalized Differential Construct
First Order (Jacobian)	$\mathcal{J}F_1(x)$	$J_2(x) \in \partial F_2(x)$
Second Order (Hessian)	$\nabla^2 f_1(x)$	$H_2(x) \in \partial\nabla f_2(x)$

Table 1. Notation for Differential Constructs

We prove several propositions using the upper semicontinuity of $\partial\nabla f$, where f is a C^1 convex function.

Proposition 1.3.1. *Neighborhood of Invertibility.*

Let f be a C^1 convex function (which implies that $\partial\nabla f(x)$ exists everywhere). Let x_* be given and assume any generalized Hessian $H(x_*) \in \partial\nabla f(x_*)$ is invertible. Then $\beta = \sup\{\|H^{-1}(x_*)\| \mid H(x_*) \in \partial\nabla f(x_*)\}$ satisfies $\beta < +\infty$. Further, there exists a neighborhood $N(x_*, \epsilon)$ of x_*, such that if $x \in N(x_*, \epsilon)$, then any $H(x) \in \partial\nabla f(x)$ is invertible and $\|H^{-1}(x)\| \leq 2\beta$.

Proof. Since the generalized Hessian set is compact, then $\beta = \max\{\|H^{-1}(x_*)\| \mid H(x_*) \in \partial\nabla f(x_*)\}$ is well-defined. Using the upper semicontinuity of the generalized Hessian map $\partial\nabla f$, choose ϵ small enough, so that for $x \in N(x_*, \epsilon)$ and any $H(x) \in \partial\nabla f(x)$, then some $H(x_*)$ satisfies:

$$\|H(x) - H(x_*)\| < \frac{1}{2\beta}.$$

Then, for any $x \in N(x_*, \epsilon)$, any $H(x) \in \partial\nabla f(x)$,

$$\|H^{-1}(x_*)(H(x) - H(x_*))\| \leq \|H^{-1}(x_*)\|\,\|H(x) - H(x_*)\| \leq \beta\frac{1}{2\beta} = \frac{1}{2}.$$

By the theorem of invertibility of perturbed matrices, the matrix $H(x)$ is invertible and satisfies:

$$\|H^{-1}(x)\| \leq \frac{\|H^{-1}(x_*)\|}{1 - \|H^{-1}(x_*)(H(x) - H(x_*))\|} \leq 2\beta.$$

||

If f is C^1 convex, using the definition of the generalized Hessian and the property that ∇^2 is symmetric, positive semi-definite for convex functions, then any generalized Hessian of a C^1 convex functional is symmetric, positive semi-definite, being a convex combination of a finite number of limits of positive semi-definite matrices. See Hiriart-Urruty[1986].

Proposition 1.3.2. *Neighborhood of Positive Definiteness.*

Let f be a C^1 convex function. Let x_* be given. Assume that any generalized Hessian $H(x_*) \in \partial\nabla f(x_*)$ is positive definite. Then there exists a neighborhood $N(x_*, \epsilon)$ of x_*, such that if $x \in N(x_*, \epsilon)$, any generalized Hessian in $H(x) \in \partial\nabla f(x)$ is positive definite.

Proof. Since the generalized Hessian set is compact, one can solve for $H(x_*) \in \partial\nabla f(x_*)$ and d where $\|d\| = 1$:

$$\lambda := \quad \min \quad \langle d, H(x_*)d\rangle > 0$$

$$\text{subject to: } \|d\| = 1$$

$$H(x_*) \in \partial\nabla f(x_*).$$

Using the upper semicontinuity of the map $\partial\nabla f(x_*)$, choose $\epsilon > 0$ small enough so that for $x \in N(x_*, \epsilon)$, any $H(x) \in \partial\nabla f(x)$, then some $H(x_*) \in \partial\nabla f(x_*)$ satisfies:

$$\|H(x_*) - H(x)\| \leq \frac{\lambda}{2}.$$

Then, for any d where $\|d\| = 1$, by the Cauchy-Schwartz inequality,

$$\lambda + \langle d, -H(x)d\rangle \leq |\langle d, (-H(x) + H(x_*))d\rangle| \leq \|H(x) - H(x_*)\| \leq \frac{\lambda}{2},$$

which implies that $\langle d, H(x)d\rangle \geq \frac{\lambda}{2}$. ||

1.4 Han's Algorithm for Solving Linear Inequalities

Han [1980] describes a method to solve a linear system of inequalities in a least square sense. We give a brief description of this method, and develop and verify properties of the iterates.

Let A be a linear transformation from $\Re^n$ to $\Re^m$, and let $b \in \Re^m$. Han's method provides a solution to a system of inequalities, $Ax \leq b$, whether feasible or infeasible, by solving in a least square sense:

$$\min f(x) := \frac{1}{2}\langle (Ax - b)_+, (Ax - b)_+\rangle, \tag{1.4.1}$$

where $(Ax - b)_+$ denotes the projection of $Ax - b$ onto $\Re^m_+$. The optimal residual $z_* = (Ax_* - b)_+$, where x_* solves (1.4.1), is unique and satisfies $A^T z_* = 0$. One can represent (1.4.1) as a convex quadratic programming problem with linear constraints in $\Re^{m+n}$:

$$\min_{x,z} \quad \frac{1}{2}\langle z, z\rangle$$

$$\text{subject to:} \tag{1.4.2}$$

$$Ax - b \leq z$$

Since $Ax - b \leq z$ is strictly feasible, and since the objective function is convex and quadratic, there exists a solution. The Karusch-Kuhn-Tucker (KKT) conditions for a solution (z, x) to (1.4.2) are:

$$A^t z = 0,$$

$$Ax - b \leq z,$$

$$z \geq 0,$$

$$\langle z, Ax - b - z\rangle = 0.$$

Denote $I(x) := \{i | \langle a_i, x - b_i \rangle \geq 0\}$. In most cases, the dependence of x is evident in context, so for simplicity of notation, $I = I(x)$ is used. Denote A_I as the submatrix of A containing rows $i \in I$. Notice that if $I(x) = \emptyset$, then x is a solution. Given an iterate x_k, let $I = I(x_k)$ and define the search direction d_k as:

$$d_k := -A_I^{\dagger}(A_I x_k - b_I), \qquad (a1)$$

where $A_I^{\dagger}$ is the psuedo-inverse. $A_I^{\dagger}$ gives the unique least norm least square projection of $-(A_I x_k - b_I)$ onto the subspace spanned by $a_i, i \in I = I(x_k)$. Han verifies that:

$$A_I^T A_I d_k = -A^T(Ax_k - b)_+ = -\nabla f(x_k).$$

Calculate the next iterate, $x_{k+1} := x_k + t_k d_k$ where λ_k solves the exact line search:

$$\min_{\lambda} \quad \frac{1}{2} f(x_k + \lambda d_k). \qquad (a2)$$

Han shows finitely terminating global convergence to a global (but possibly, non-unique) minimum. One can express the algorithm succinctly, as follows: Let the iterate x_k be given. Let d_k be calculated as the unique least norm solution of least square solution of:

$$A_I d_k = -(A_I x_k - b_I). \qquad (b1)$$

Let $z_k := (A_I x_k - b_I)$, and let $\lambda_k \geq 0$ be calculated so that:

$$\langle (z_k + \lambda_k A d_k)_+, A d_k \rangle = 0. \qquad (b2)$$

Because the line search function is piece-wise quadratic and convex, the exact minimizer λ_k is easily found. Let $x_{k+1} := x_k + \lambda_k d_k$. If $\nabla f(x_{k=1}) = 0$, then x_{k+1} solves (1.4.1); otherwise, iterate.

Properties

We show that the iterates satisfy several other properties in the following.

1. For $z_{k+1} := z_k + \lambda_k A d_k$, since the objective function is convex, the subgradient relationship implies the objective function decrease satisfies:

$$\langle z_k^+, z_k^+ \rangle - \langle z_{k+1}^+, z_{k+1}^+ \rangle \geq 2\lambda_k \| A_I d_k, A_I d_k \|^2$$

2. $\langle z_k^+, z_k^+ \rangle - \langle r_{k+1}^+, r_{k+1}^+ \rangle = \langle z_k^+, z_k \rangle - \langle z_{k+1}^+, z_k + \lambda_k A d_k \rangle = \langle z_k^+ - z_{k+1}^+, z_k \rangle$, by property (b2) and properties of the projection, z^+.
3. If $I_k = I_{k+1}$, then $\langle z_k^+, z_k^+ \rangle - \langle z_{k+1}^+, z_{k+1}^+ \rangle = \langle z_I, -\lambda_k A_I d_k \rangle$, by computation.
4. If $A_I d_k = 0$, then x_k is a solution, that is, $A^T z_k^+ = 0$.
5. If $I_k = I_{k+1}$, then $\langle z_k^+, z_k^+ \rangle - \langle z_{k+1}^+, z_{k+1}^+ \rangle = 0$. Letting $I = I_k$ and from 1 above, given that

$$\langle z_k^+, z_k^+ \rangle - \langle z_{k+1}^+, z_{k+1}^+ \rangle \geq 2\lambda_k \| A_I d_k, A_I d_k \|^2,$$

   then

$$\begin{aligned} \langle z_I, -\lambda A_I d_k \rangle - 2\langle \lambda A_I d_k, A_I d_k \rangle &\geq 0, \\ \langle z_I + \lambda A_I d_k, -\lambda A_I d_k \rangle - \langle \lambda A_I d_k, A_I d_k \rangle &\geq 0, \\ -\langle \lambda A_I d_k, A_I d_k \rangle &\geq 0, \end{aligned} \tag{1.4.3}$$

   which implies that $A_I d_k = 0$. This means that, if at any time the index set of the current iterate I_k is the same as the index set of the next iterate I_{k+1}, then the algorithm converges at step $k + 1$.
6. By the Projection Theorem, if d_k solves (a1), then for all $p \in R^n$,

$$\langle z_I + A_I d_k, A_I p \rangle = 0,$$

   which implies:

$$\langle z_I + A_I d_k, A p \rangle = 0.$$

7. By (b1) and (b2), there is a special conjugacy relationship of the current direction d_k, and the new direction d_{k+1}:

$$\langle A d_k, A_{I_{k+1}} d_{k+1} \rangle = 0.$$

Han's algorithm has, in practice, fast convergence to a solution, whether feasible or infeasible. Han shows that after a finite number of iterations, the iterates satisfy a relationship:

$$I_k \subset I_{k+1}, \tag{1.4.4}$$

and from (1.4.3), this implies that after at most m more iterations, then the algorithm must converge. This special neighborhood of a solution x_* is called an *identification neighborhood*, in that unmet and active constraints of the solution are a superset of those neighboring points' unmet and active constraints.

1.5 Background on Least Square Algorithms for the Smooth Case

In addition to Han's method for solving a system of inequalities in a least square sense, there are other minimization methods available for solving a system of equations in a least square sense, a smooth problem. For example, if one wishes to solve for

$$F(x) = 0,$$

where $F : \Re^n \mapsto \Re^m$ is C^2, one defines

$$f(x) := \frac{1}{2}\langle F(x), F(x)\rangle \tag{1.5.1}$$

and solves:

$$\inf f(x).$$

These minimization algorithms depend on two important calculations at each iterate x, deriving a search direction, d, and a step size λ, such that the step taken is λd. Most minimization algorithms for least squares find a descent direction by solving $Bd = -\nabla f(x)$, for d, where $\nabla f(x)$ is the gradient of the current iterate and B is an $n \times n$ matrix determined by the algorithm. Denote $J(x) := \mathcal{J}F(x)$, the Jacobian of F at the current iterate x.

1) If $B = I$, where I is the identity matrix, then d is the steepest descent direction.
2) If $B = J^T(x)J(x)$, then d is the Gauss-Newton direction.
3) If $B = J^T(x)J(x) + \mu I$, where $\mu \geq 0$ is chosen to assure that $J^T(x)J(x) + \mu I$ is positive definite, then d is the Levenberg-Marquardt direction.
4) If B is a rank 1 or rank 2 update approximation of the Hessian of f, which satisfies the quasi-Newton condition, then d is a quasi-Newton direction.
5) If $B = H(x)$, the Hessian of f at x, then d is the Newton direction.

Notice that calculating the search direction requires that the function f be at least once differentiable, as in the case of the steepest descent direction. Newton's method uses second order information. An important property of B is that it be positive definite to assure a descent direction. Therefore, one must define an algorithm which accounts for this. If B is not positive definite, it may not be invertible, and the Gauss-Newton and Newton's methods can be generalized by solving $Bd = -g(x)$ in a least square sense.

There are specialized algorithms to solve least square problems. Using second order methods, such as Newton's method or quasi-Newton methods to solve least square problems do not take advantage of the special structure of the least square objective function. Notice that the Hessian of the objective function can be expressed: $\nabla^2 f(x) = J^T(x)J(x) + Q(x)$, where

$$Q(x) := \sum_{i=1}^{m} F_i(x)\Gamma_i(x),$$

where $\Gamma_i(x)$ is the Hessian of $F_i(x)$. In cases where the problem has a zero or small residual solution, $Q(x)$ is dominated by $J(x)^T J(x)$ close to the solution. See Varah [1990] and Gill, Murray and Wright [1981]. A small residual problem is characterized by properties of the Jacobian of F at the solution and the magnitude of the residuals. Because of the special structure of the non-linear least squares problem, one usually uses the Gauss-Newton method or the Levenberg-Marquardt algorithm for zero or small residual problems.

Newton's method and quasi-Newton methods are also applicable for solving non-linear least squares, particularly for those having large residuals. In section 4 we adapt these methods to solve systems of inequalities, by selecting a generalized Jacobian or Hessian at each iterate.

Inexact line search methods compare the predicted step size and the decrease in function value versus the actual step size and function value. For example, given α, β, such that $0 < \alpha < \beta < 1$, the Armijo line search conditions require λ to satisfy,

$$\begin{aligned} f(x+\lambda d) &\leq f(x) + \alpha\langle g(x), \lambda d\rangle, \\ \langle g(x+\lambda d), \lambda d\rangle &\geq \beta\langle g(x), \lambda d\rangle. \end{aligned} \tag{1.5.2}$$

Conditions on the function f which assure the existence of λ to satisfy (1.5.2) are that ∇f be continuous, that the search direction d is a descent direction of f at x, and that $\{f(x+\lambda d) \mid \lambda > 0\}$ is bounded below.

The model trust region (or restricted step) method uses a model function, typically a second order (Taylor series) approximation of f at x to compare the predicted versus actual decrease in the function, adjusting the step size appropriately. Suppose $M(s) := f(x) + \langle g(x), s\rangle + \frac{1}{2}\langle s, H(x)s\rangle$, where $g(x)$ is the gradient of f at x and $H(x)$ is the Hessian of $f(x)$. Then the Taylor series approximation gives: $f(x+s) \approx M(s)$, about x. The method finds a specific step, s, to satisfy agreement between the model function and the function f. An inital $\delta > 0$ is chosen. The subproblem below is solved for s:

$$\min_s \quad M(s) \quad \text{subject to } \|s\| \leq \delta. \tag{1.5.3}$$

Calculate the ratio:

$$r := \frac{f(x) - f(x+s)}{M(0) - M(s)} = \frac{f(x) - f(x+s)}{f(x) - M(s)}. \tag{1.5.4}$$

1) If $r < .25$, then set the new $\delta := \|s\|/4$.
   If $r > .75$, and $\|s\| = \delta$, then set the new $\delta := 2\delta$,
   otherwise, set the new $\delta := \delta$.
2) If $r \leq 0$, then set new $x := x$; else, set new $x := x + s$.

The model trust region method may not take a step at each iteration, and the method does not require that the Hessian be positive definite. The choices of parameter, above, for example, .25, and .75, are based on computational experience. Solving subproblem (1.5.3) is often done approximately.

The model trust region method requires a second order approximation of the function. In section 4, we show that using the generalized Hessian provides this approximation.

2 Piece-wise Linear Assignment Problem

In this section, an assignment problem having a piece-wise linear objective cost function is described. The piece-wise linear objective function of use here is the maximum of a constant function and one or several linear functions. This problem is the motivation and an application of algorithms developed in sections 4 and 6. This type of piece-wise linear objective cost function is useful in modeling the costs of assigning jobs to individual servers, each of which have fixed minimum charges and linear rates above the minimum. The problem is expressed as a $\{0,1\}$ programming problem for which a strong relaxation is derived. The dual of the relaxed primal problem is derived, as well as optimality conditions for the relaxed primal solutions. Alternative methods for solving for $\{0,1\}$ solutions within the relaxed problem solution set are described and compared.

This assignment problem belongs to a family of problems, called the uncapacitated facility location problem. It is a generalization of the problem described in Conn and Cornuéjols[1990].

2.1 Statement of Problem

Let $i \in I = \{1, 2, \ldots, m\}$ represent jobs to be assigned to servers $j \in J = \{1, 2, 3, \ldots, n\}$. Let $1^m \in \Re^m$ represent the unit demand vector of the jobs. For each $j \in J$, $i \in I$, let $y_{i,j}$ be the relative amount of demand i, assigned to server j, i.e., $y \in R^{mn}$. The variable y is constrained so that $0 \le y_{i,j}$, for each $j \in J$, $i \in I$. Further, the variable y must meet the demand for each i. That is, for all $i \in I$,

$$\sum_{j \in J} y_{i,j} = 1.$$

Let positive linear cost coefficients, $c_{i,j}$, and positive minimum cost constants, ϕ_j, be associated with each server j and job i. That is, assume for all $i \in I$ and $j \in J$, that $c_{i,j} > 0$ and $\phi_j > 0$. To simplify notation, interpret the vector $y_{*,j}$ to be the elements of $y \in \Re^{mn}$, associated with server j. The cost function for each server j is:

$$\begin{cases} \max(\phi_j, \langle c_{*,j}, y_{*,j} \rangle) & \text{if any } y_{i,j} > 0, \\ 0 & \text{if } y_{i,j} = 0, \forall i \in I. \end{cases}$$

This server cost function represents a flat minimum charge with linear rates beyond the minimum. Some restaurants, for example, charge a flat table charge, even if the sum of the menu items fall beneath this price. Figure 1 shows the graph of an example server cost function.

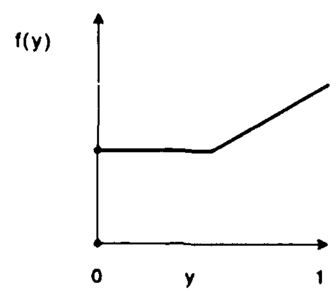

Figure 1. Example Server Cost Function

One can form the strong relaxations of the cost functions by introducing variables, $1 \geq w_j \geq 0$, for $j \in J$, and adding the constraints that

$$1 \geq w_j \geq y_{i,j}, \forall i, j.$$

The relaxed primal problem for the assignment problem is then

$$\inf_{w,y} \quad \sum_j f_j(w_j, y_{*,j}) \tag{2.1.1}$$

$$\text{subject to: } \forall i \in I \qquad \sum_{j \in J} y_{i,j} = 1,$$

where $f_j(w_j, y_{*,j}) := \max(w_j \phi_j, \langle c_{*,j}, y_{*,j} \rangle) + \delta(w_j, y_{*,j} \,|\, 0 \leq \ y_{i,j} \leq w_j \leq 1, \forall i \in I)$. Here, δ is the indicator function; that is,

$$\delta(w_j, y_{*,j}) := \begin{cases} 0 & \text{if the condition is satisfied,} \\ +\infty & \text{if the condition is not satisfied.} \end{cases}$$

Notice that f_j is convex, being the pointwise maximum of a constant function and a linear function. By adding a variable $z \in \Re^n$, the problem can be expressed in standard form with $mn + 2n$ variables and $2mn + m + 3n$ constraints as follows:

$$(P): \qquad \min \sum_j z_j$$

$$\begin{aligned} &\text{subject to:} \\ w_j - y_{i,j} &\geq 0, \quad \forall i, j \\ -\phi_j w_j + z_j &\geq 0, \quad \forall j \\ -\sum_i c_{i,j} y_{i,j} + z_j &\geq 0, \quad \forall j \\ \sum_j y_{i,j} &= 1, \quad \forall i \\ w_j &\leq 1, \quad \forall j \\ y_{i,j} &\geq 0, \quad \forall i, j. \end{aligned} \tag{2.1.2}$$

The relaxed primal problem (P) is feasible, and the optimal value is finite, since the cost coefficients are all positive and the equality constraint on y assures a closed bounded feasible region for the y variable. Thus, by the Duality Theorem of Linear Programming, there exists solutions to the dual of the relaxed problem. See, for example, Kolman and Beck[1980].

2.2 Dual of the Relaxed Problem

Define dual variables, $v \in \Re^{mn}$, $r, s, t \in \Re^{n}$, and $u \in R^{m}$, so that the dual, directly derivable from (P) is :

$$
\begin{aligned}
(D): \qquad \max \quad & \sum_i u_i - \sum_j t_j \\
\text{subject to:} & \\
\sum_i v_{i,j} - \phi_j s_j &= 0, \quad \forall j \\
-v_{i,j} + u_i - c_{i,j} r_j &\leq 0, \quad \forall i, j \\
s_j + r_j &= 1, \quad \forall j \\
v_{i,j} &\geq 0, \quad \forall i, j \\
t_j &\geq 0, \quad \forall j \\
s_j &\geq 0, \quad \forall j \\
r_j &\geq 0, \quad \forall j \\
u_i & \text{ unconstrained}, \ \forall i.
\end{aligned}
\tag{2.2.1}
$$

By substituting

$$s_j = \left(\sum_i v_{i,j} \right) / \phi_j \geq 0,$$

since $v_{i,j} \geq 0$, and $\phi_j > 0$; and substituting, $r_j = 1 - s_j \geq 0$, one can eliminate the variables r and s to obtain:

$$
\begin{aligned}
(D): \qquad \max \quad & \sum_i u_i - \sum_j t_j \\
\text{subject to:} & \\
-v_{i,j} + u_i - c_{i,j} \left(1 - \phi_j^{-1} \sum_k v_{k,j} \right) &\leq 0, \quad \forall i, j \\
\sum_i v_{i,j} &\leq \phi_j, \quad \forall j \\
t_j &\geq 0, \quad \forall j \\
v_{i,j} &\geq 0, \quad \forall i, j \\
u_i & \text{ unconstrained}, \ \forall i.
\end{aligned}
\tag{2.2.2}
$$

The dual has $mn + m + n$ variables and $2mn + 2n$ constraints. Solving the dual problem may be preferable, if there are fewer variables and constraints. Then, after solving the dual, one must retrieve a solution.

2.3 Solving for the Primal from the Relaxed Dual

Since (P) and (D) are linear programming duals of each other, and if u, v, t (and implicitly r and s) solve (D), then there exists w, y, z which solve (P). The optimality conditions which must be satisfied can be expressed, for optimal u, v, t and

$$s_j = \left(\sum_i v_{i,j}\right) / \phi_j,$$

$$r_j = 1 - s_j.$$

Then, from (P), it follows:

$$\begin{aligned}
w_j - y_{i,j} &\geq 0, \quad \forall i, j \\
-\phi_j w_j + z_j &\geq 0, \quad \forall j \\
-\sum_i c_{i,j} y_{i,j} + z_j &\geq 0, \quad \forall j \\
\sum_j y_{i,j} &= 1, \quad \forall i \\
w_j &\leq 1, \quad \forall j \\
y_{i,j} &\geq 0, \quad \forall i, j.
\end{aligned}$$

Further, by complementarity conditions for (P) and (D), the following must be satisfied:

$$\begin{aligned}
v_{i,j}(w_j - y_{i,j}) &= 0, \quad \forall i, j \\
s_j(-\phi_j w_j + z_j) &= 0, \quad \forall j \\
r_j(-\sum_i c_{i,j} y_{i,j} + z_j) &= 0, \quad \forall j \\
t_j(w_j - 1) &= 0, \quad \forall j \\
y_{i,j}(-v_{i,j} + u_i - c_{i,j} r_j) &= 0, \quad \forall i, j.
\end{aligned}$$

Using the above conditions, one can express the solution for w, y, z as a solution to a system of linear inequalities with respect to the following index sets:

$$\begin{aligned}
J^{=} &:= \{j \mid r_j \in (0,1)\} &&\iff w_j \phi_j = \sum_i c_{i,j} y_{i,j} = z_j \\
J^{\geq} &:= \{j \mid r_j = 1\} &&\iff \sum_i c_{i,j} y_{i,j} = z_j \geq w_j \phi_j \\
J^{\leq} &:= \{j \mid r_j = 0\} &&\iff \sum_i c_{i,j} y_{i,j} \leq w_j \phi_j = z_j \\
I_j^0 &:= \{i \mid -v_{i,j} + u_i - c_{i,j} r_j < 0\} &&\iff y_{i,j} = 0 \\
I_j^w &:= \{i \mid v_{i,j} > 0\} &&\iff y_{i,j} = w_j.
\end{aligned}$$

Solve the following system for w, y, and z: (Note that z is an unnecessary variable.)

$$
\begin{aligned}
y_{i,j} &= 0, \quad \forall i \in I_j^0, \forall j \\
y_{i,j} &\geq 0, \quad \forall i, j \\
\sum_j y_{i,j} &= 1, \quad \forall i \\
y_{i,j} &= w_j, \quad \forall i \in I_j^w, \forall j \\
w_j &\leq 1, \quad \forall j \\
y_{i,j} &\leq w_j, \quad \forall i, j \\
t_j(w_j - 1) &= 0, \quad \forall j \\
\sum_i c_{i,j} y_{i,j} &= z_j = w_j \phi_j, \quad \forall j \in J^= \\
\sum_i c_{i,j} y_{i,j} &= z_j \geq w_j \phi_j, \quad \forall j \in J^\geq \\
\sum_i c_{i,j} y_{i,j} &\leq w_j \phi_j = z_j, \quad \forall j \in J^\leq.
\end{aligned}
\tag{2.3.1}
$$

2.4 Sample Problem

In order to illustrate the primal and dual relationship, the following simple problem is shown. Let $m = 4$, $n = 3$; let $\phi = (26, 30, 30)$; let

$$
c = \begin{pmatrix} 20 & 25 & 25 \\ 6 & 5 & 5 \\ 40 & 40 & 44 \\ 30 & 30 & 22 \end{pmatrix}.
$$

The optimal value is $87\frac{3}{11}$, achieved at, for example, $t = (1, 0, 1)$ and $w = (1, 0, 1)$ and:

$$
u = \begin{pmatrix} 20 \\ 5 \\ 40 \\ 22\frac{3}{11} \end{pmatrix}, \quad v = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & \frac{5}{11} \\ 0 & 0 & 0 \\ 0 & 0 & 2\frac{3}{11} \end{pmatrix}, \quad y = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ \frac{41}{44} & 0 & \frac{3}{44} \\ 0 & 0 & 1 \end{pmatrix}.
$$

$s_1 = s_2 = 0$, and $s_3 = \frac{1}{11}$. Thus, $r_1 = r_2 = 1$, and $r_3 = \frac{10}{11}$. We show that this solution satisfies the optimality conditions (2.3.1).

First, consider $j = 2$, where $y_{i,2} = 0$ for all i, satisfying $I_2^0 = \{1, 4\}$, and $I_2^w = \emptyset$. Further, $w_2 = 0$ is consistent, since $J^\geq = \{1, 2\}$ and

$$
\langle c_{*,2}, y_{*,2} \rangle = 0 \geq w_2 \phi_2 = 0 * 30 = 0.
$$

Second, consider $j = 3$. $I_3^0 = \{1\}$ and $I_3^w = \{2, 4\}$. It must be that $r_3 = \frac{10}{11} \in (0, 1)$ and $J^= = \{3\}$. It must be that $y_{4,3} = 1$, which implies that $w_3 = 1 = y_{2,3}$. To satisfy,

$$
\langle c_{*,3}, y_{*,3} \rangle = w_3 \phi_3 = 30,
$$

then $y_{3,3} = \frac{3}{44}$.

Lastly, consider $j = 1$, where $I_1^0 = \{2, 4\}$ and $I_1^w = \emptyset$. Since $1 \in J^{\geq}$, then

$$\langle c_{*,1}, y_{*,1} \rangle = 57\frac{3}{11} > w_1\phi_1 = 26.$$

$y_{1,1} = 1 \leq w_1 = 1$; $y_{1,3} = \frac{41}{44} \leq w_1$.

For $i = 3$, $\sum_j y_{3,j} = 1$. Further, for all i, $\sum_j y_{i,j} = 1$.

The full set of solutions of this problem include: $y_{i,j} = 0$, except as follows:

$$y_{1,1} = w_1 = 1, \quad y_{2,3} = y_{4,3} = w_3 = 1, \quad y_{3,3} = \frac{3}{44},$$

and $w_2, y_{3,2}$, and $y_{3,1}$ satisfy:

$$w_2 \leq 1,$$
$$y_{3,1} \geq \frac{6}{40},$$
$$y_{3,2} = \frac{41}{44} - y_{3,1},$$
$$0 \leq y_{3,2} \leq w_2 \leq \frac{4}{3}y_{3,2}.$$

See Figure 2, Example Solution Set in 2 Dimensions, w_2 and $y_{3,1}$.

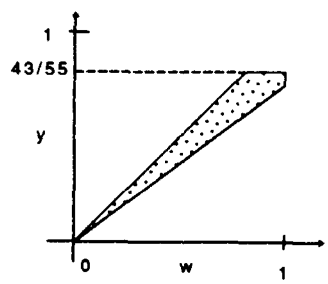

Figure 2. Example Solution Set in 2 Dimensions

2.5 Methods to Solve Assignment Problem

The attractiveness of the strong relaxation of the facility location problem is seen in experimental results. For example, Cornuéjols et al. [1989] define the convex hull of the uncapacitated facility problem solutions as the "uncapacitated facility polytope". They point out that often there is a $\{0,1\}$ solution within the relaxed solution set; and if not, a solution can be found after a small number of cutting plane iterations. Their experimental results show a small difference in the objective function values between the two problems, the relaxed and $\{0,1\}$ problem. Thus, most methods to solve this problem use linear programming relaxations, although there are great variations on how to solve these subproblems. Traditional, primal simplex and cutting plane methods are used. Conn and Cornuéjols [1990], for their problem, solve the dual of the relaxed primal using projections. If solving the dual is more efficient, one advantage of this may be that a (perhaps, incomplete) set of primal solutions are characterized through the complementary slackness conditions of each dual solution. It is desirable to find an efficient method to search the set of primal solutions generated from a given dual solution.

The primal solution set generated by a given dual solution is a polyhedral convex set being the solution of (2.3.1). The conditions on w assure that w is bounded. The condition on the y variable, $\sum_j y_{i,j} = 1$, for all i, and that $y_{i,j} \geq 0$, for all i,j assure that y is bounded. So the solution set of the relaxed problem is also a polytope. Searching for $\{0,1\}$ solutions among the vertices of this polytope, via direct search, could require 2^n iterations, where at each iteration, a solution in the variable y must be computed or found not to exist.

The approach proposed in this dissertation is to apply a penalty-type method, adding equations to the system, which can be solved using continuous methods. As an example of a penalty for $\{0,1\}$ conditions on w, one may choose to add the condition, $w_j(1-w_j) = 0$, for all j. This non-linear condition has the property that, if met, guarantees a $\{0,1\}$ solution in w_j. However, the non-linear property demands more powerful solution methods.

One can express the system of equalities and inequalities in (2.3.1), adding the condition, $w_j(1-w_j) = 0$ in more convenient notation. Use a single variable x to represent the variable (w,y), noting that z is an auxillary variable determined by the variable (w,y). Define the function F to represent the nonlinear coercive function of w. For example, $F_j : \Re \mapsto \Re$, and $F_j(w_j) := w_j(1-w_j)$. Let the matrix A and the vector b represent the linear transformations in the inequalities, and the matrix C and the vector d represent the linear transformations in the equalities. Then, solve the system below, finding x which satisifies:

$$
\begin{aligned}
F(x) &= 0, \\
Ax &\leq b, \\
Cx &= d.
\end{aligned}
\tag{2.5.1}
$$

This system has linear inequalities and nonlinear and linear equalities. In the introduction, Han's method to solve systems of linear inequalities in a least square sense was described. One approach to solve (2.5.1) is to extend Han's method to include non-linear equations. Burke [1984] describes one algorithm to solve the linear inequalities with non-linear equations in a least square sense. The search direction is derived by projecting (least norm least square) the negative of the residuals onto the space spanned by the coefficients of unmet inequality constraints. He uses an Armijo line search to gain global convergence. The least square least norm projection uses a singular value decomposition to find a search direction. We will show that common search direction methods are applicable, such as Gauss-Newton, and quasi-Newton methods. These methods, using direct inverses or projections using QR decomposition with column pivoting, are finite processes for finding a search direction and can be more efficient in computation. Further, the other major global convergence technique, the trust region method, is examined, considering various model functions.

It is not advisable to use a penalty function having a large penalty parameter or a barrier function to coerce a solution. Large penalty parameters are known to give ill-conditioned subproblems and are computationally impractical, especially for this problem, since solutions will necessarily occur at the boundary of the feasible set.

Using Newton's method to directly solve the system of equations:

$$\begin{aligned} F(x) &= 0, \\ (Ax - b)_+ &= 0, \\ Cx &= d. \end{aligned} \tag{2.5.2}$$

is difficult due to the non-differentiability property of $(Ax - b)_+$. Advances in non-differentiable optimization have yielded new methods; but, to-date, these "bundle " methods have suffered from poor stability and rates of convergence. See Lemarechal[1986]. Recent developments, Schramm and Zowe[1991], use trust-region features to increase the bundle method performance. However, the special structure of this problem suggests again that more efficient methods can be found.

3 Properties of the Least Square Formulation for Solving a System of Linear Inequalities

From the discussion at the end of section 2.5, a more promising approach to solve (2.5.1) is a least square approach using a small penalty parameter, such as 1. This approach has two strong properties. First, the objective function is differentiable. Secondly, we show in this section that the least squares problem with small penalty parameter is numerically stable under small perturbation due to inaccurate computation. Further, there are many choices of algorithms available for minimization of differentiable functions, several of which will be analyzed and adapted for use in this problem. The system is solved by minimizing the norm of the residuals of the inequalities and equalities.

In section 3.1, we define the least square formulation for solving a system of nonlinear equations and linear equalities and in section 3.2 develop numerical stability properties of the least square objective function. In section 3.3, a special neighborhood of a designated point, called an identification neighborhood is defined in relation to a finite system of weak linear inequalities. Points in an identification neighborhood of the designated point satisfy the same strict linear inequalities as does the designated point, with the exception of those weak linear inequalities which the designated point satisfies with equality. This means that the points in an identification neighborhood are on the same side of all separating hyperplanes (determined by the weak linear inequalities) as the designated point except for those hyperplanes on which the designated point lies. Further, in section 3.3, generalized differential constructs, first order and second order, are defined for the system of inequalities and the least square objective function. Properties of Taylor series expansion based on generalized differential constructs to locally approximate the least square objective function are analyzed.

3.1 Statement of Least Square Problem

We wish to find $x \in R^n$ which solves the following system of linear inequalities and equalities and non-linear equalities:

$$\begin{aligned} F_1(x) &= 0, \\ Ax &\leq b, \end{aligned} \tag{3.1.1}$$

where

$$F_1 : \Re^n \mapsto \Re^{m1} \text{ and } F_1 \text{ is non-linear,}$$

$$A : \Re^n \mapsto \Re^{m2} \text{ and } A \text{ is linear.}$$

This is a mixed system of linear inequalities and equalities and non-linear equalities. Notice that the linear equations and the non-linear equations are combined into a single multivariate function, F_1. Theoretically, these are treated together; but one should exploit the simple linear equations properties in computational algorithms. One can transform this problem into a least squares minimization problem. Denote a_i^T to be the ith row of A and a_i^j to be the jth element of the ith row of A. Recall that $(Ax - b)_+$ represents the projection of $Ax - b$ onto $\Re^{m2}_+$. We define $F : \Re^n \mapsto \Re^{m1+m2}$

$$F(x) := \begin{pmatrix} F_1(x) \\ (Ax - b)_+ \end{pmatrix} = \begin{pmatrix} F_1(x) \\ F_2(x) \end{pmatrix}.$$

Let $f : \Re^n \mapsto \Re$ be defined:

$$f(x) := f_1(x) + f_2(x), \text{ where}$$
$$f_1(x) := \frac{1}{2}\langle F_1(x), F_1(x)\rangle,$$
$$f_2(x) := \frac{1}{2}\langle (Ax-b)_+, (Ax-b)_+\rangle.$$

Then, instead of solving (3.1.1), find x which solves

$$\inf \quad f(x). \tag{3.1.2}$$

Although the two problems, (3.1.1) and (3.1.2), are not equivalent, notice that if x_* solves $\inf f(x)$ in (3.1.2), then x_* solves (3.1.1). If $\inf f(x) > 0$ in (3.1.2), then (3.1.1) has no solution. The least square algorithms do not guarantee that a global infimum is returned to solve (3.1.2). They provide a local minimum stationary point x_* of (3.1.2) which may or may not be a solution for (3.1.2). For example, if f is not convex, there may be non-global, local minima. f given by (2.5.1) is not convex. Even though the algorithm may return a local minimum stationary point, we show in section 7 that a local minimum solution of (3.1.2) provides important information for solving the assignment problem.

3.2 Stability of the Least Square Problem for Solving Systems of Linear Inequalities

One aspect of numerical stability that is related to solving optimization problems is the sensitivity of the solution to the problem due to small perturbations or inaccuracies of computation. One, intuitively, would characterize numerically stable problems as ones which, given small inaccuracies in computation, would give small inaccuracies in the solution. Certain problem formulations are known to be numerically unstable, such as penalty functions having large penalty parameters and barrier functions, which have solutions on the boundary of the feasible set. In this section, we show that the formulation of the least square problem for solving systems of inequalities is numerically stable to small inaccuracies in computation. See Robinson [1975,1976] where he develops stability theory for systems of inequalities, both the linear and non-linear case.

Consider the linear transformation A in (3.1.1) where $A : \Re^n \mapsto \Re^{m2}$, a right hand side $b \in \Re^{m2}$, and the variable $x \in \Re^n$. The system of linear inequalities is:

$$Ax \leq b. \tag{3.2.1}$$

Define $\Omega := \{x | Ax \leq b\}$ as the solution set of (3.2.1), where Ω may be empty. We say that (3.2.1) is solvable if Ω is nonempty. Robinson defines the nonempty solution set Ω as *stable* if for each $x_0 \in \Omega$ there are some positive constants β and δ, such that for any linear transformation $A' : \Re^n \mapsto \Re^m$, and $b' \in \Re^m$, where $\|A - A'\| + \|b - b'\| \leq \delta$, the distance from x_0 to the solution set of the perturbed system, $\Omega' := \{x | A'x \leq b'\}$ satisfies

$$\text{dist } (x_0, \Omega') \leq \beta\rho(x_0), \tag{3.2.2}$$

where

$$\rho(x) := \inf_{k \in \Re^m_+} \|b' - A'x - k\|. \tag{3.2.3}$$

Notice the $\rho(x) = \|(A'x - b')_+\|$. It is not sufficient for the system (3.2.1) to be solvable to guarantee Ω is stable. We define dist (x_0, Ω') to be infinite, if Ω' is empty. A simple example of an unstable solvable system, in a one-dimensional variable x, is:

$$x \leq 0,$$

$$x \geq 0.$$

For this system, $\Omega = \{0\}$. Yet, for any $\delta > 0$, the perturbed sytem,

$$x \leq -\frac{\delta}{2},$$
$$x \geq \frac{\delta}{2},$$

has empty solution set, meaning that the solution set is unstable.

Robinson defines the system (3.2.1) to be *regular* if there exists some x such that $Ax < b$. This is a strict consistency condition. Robinson shows that a system (3.2.1) is regular if and only if the solution set Ω is stable.

Rather than formulating the problem as in (3.2.1), one can represent the problem as a least square formulation by finding $x \in \Re^n$ and $z \in \Re^{m2}$ which solve:

$$\begin{aligned} &\inf_{x,z} \ \frac{1}{2}\langle z, z\rangle) \\ &\text{subject to: } Ax - b \leq z. \end{aligned} \tag{3.2.4}$$

Recalling the KKT conditions for this problem,

$$\begin{aligned} A^T z &= 0, \\ z &\geq 0, \\ Ax - b - z &\leq 0, \\ \langle z, Ax - b - z\rangle &= 0. \end{aligned} \tag{3.2.5}$$

The system of inequalities represented in (3.2.5) include:

$$\begin{aligned} -z &\leq 0, \\ Ax - b - z &\leq 0. \end{aligned} \tag{3.2.6}$$

By extending the original problem (3.2.1) to (3.2.4), a least square formulation, the system of linear inequalities (3.2.6) is regular, and hence the solution set is stable. This is assured even if the original system (3.2.1) is not solvable, since (3.2.6) satisfies a strict consistency condition for any choice of A and b.

3.3 Approximating Properties of Taylor Series Using Generalized Differential Constructs

Assume that f_1 is twice continuously differentiable, which is the case for the subproblem (2.5.1) obtained in the assignment problem. Denote $H_1(x) := \nabla^2 f_1(x)$ to be the Hessian of f_1, and denote $J_1(x) := \mathcal{J} F_1(x)$ to be the Jacobian of $F_1(x)$. Refer to the differential notation that was outlined in Table 1 in section 1.3. This table clarifies the distinction in notation between generalized and single-valued differentials, and between differential constructs for f and F.

Proposition 3.3.1.

$(Ax-b)_+$ is uniformly Lipschitz of rank $\|A\|$ of order 1 on all of $\Re^n$. That is, for all x, y, $\|(Ax-b)_+ - (Ay-b)_+\| \leq \|A\| \, \|x-y\|$.

Proof. : For all i,

$$|(\langle a_i, x\rangle - b_i)_+ - (\langle a_i, y\rangle - b_i)_+| = \begin{cases} 0 & \text{if } (\langle a_i, x\rangle - b_i)_+ = (\langle a_i, y\rangle - b_i)_+ = 0, \\ |\langle a_i, x-y\rangle| & \text{if both } \langle a_i, x\rangle \geq b_i \text{ and } \langle a_i, y\rangle \geq b_i \ , \\ |\langle a_i, \lambda(x-y)\rangle| & \text{some } \lambda \in (0,1) \ , \text{ if } \langle a_i, x\rangle < b_i \text{ and } \langle a_i, y\rangle > b_i, \\ |\langle a_i, \lambda(x-y)\rangle| & \text{some } \lambda \in (0,1), \text{ if } \langle a_i, x\rangle > b_i \text{ and } \langle a_i, y\rangle < b_i. \end{cases} \tag{3.3.1}$$

The first two cases are clear. The third case, where $\langle a_i, x\rangle < b_i$ and $\langle a_i, y\rangle > b_i$, means there exists $\lambda \in (0,1)$, such that $z = y + \lambda(x-y)$ and $\langle a_i, z\rangle = b_i$. Then

$$|(\langle a_i, x\rangle - b_i)_+ - (\langle a_i, y\rangle - b_i)_+| =$$

$$|\langle a_i, z\rangle - b_i - (\langle a_i, y\rangle - b_i)| =$$

$$|\langle a_i, z-y\rangle| = |\langle a_i, \lambda(x-y)\rangle|.$$

The fourth case is shown similarly. ||

Thus the *Clarke generalized Jacobian set* of $(Ax-b)_+$, denoted $\partial(Ax-b)_+$, is well-defined. Let Ω be the set of points where the Jacobian of $(Ax-b)_+$ does **not** exist and denote $\mathcal{J}_2(y_l)$ to be the Jacobian of $(Ay_k - b)_+$, where the Jacobian does exist. Then $\partial(Ax-b)_+$ is the set of all $m \times n$ matrices $J_2(x)$, where any $J_2(x)$ is a convex combination of a finite number of limits, J_2^k, $k = 1, \ldots, K$, some K. Each limit J_2^k is obtained as the limit of a sequence of the form $\mathcal{J}_2(y_l)$, where $y_l \to x$ and $y_l \notin \Omega$. That is,

$$J_2^k := \lim_{y_l \to x} \mathcal{J}_2(y_l), \text{ where } y_l \notin \Omega.$$

Proposition 3.3.2.

$J_2(x)$ is a generalized Jacobian of $(Ax-b)_+$ if and only if

$$J_2(i,j) = \begin{cases} 0 & \text{if } \langle a_i, x\rangle - b_i < 0, \\ a_i^j & \text{if } \langle a_i, x\rangle - b_i > 0, \\ \lambda_i a_i^j & \text{if } \langle a_i, x\rangle - b_i = 0, \text{ for some } \lambda_i \in [0,1] \ . \end{cases} \tag{3.3.2}$$

Proof. For this function $(Ax-b)_+$, $\Omega = \{x | \, \langle a_i, x\rangle = b_i, \text{ for some } i\}$. By direct calculation of limits, J_2^k, one finds:

$$J_2^k = E_k A = \lim_{y_l \to x} \mathcal{J}_2(y_l), \text{ where } y_l \notin \Omega,$$

where E_k is an $m2 \times m2$ diagonal matrix, with elements e_i^k on the diagonal, where

$$e_i^k = \begin{cases} 0 & \text{if } i \in I^-(x), \\ 1 & \text{if } i \in I^+(x), \\ 0 \text{ or } 1 & \text{if } i \in I^0(x). \end{cases}$$

Then $J_2(x) \in \partial(Ax-b)_+$ means

$$J_2(x) = \sum_{k=1}^{K} \lambda_k E_k A, \text{ such that, } \sum_{k=1}^{K} \lambda_k = 1, \lambda_k \in [0,1].$$

Therefore, one can express $J_2(x) = DA$ where D is a diagonal $m2 \times m2$ matrix, with elements d_i on the diagonal, where

$$d_i = \begin{cases} 0 & \text{if } i \in I^-(x), \\ 1 & \text{if } i \in I^+(x), \\ \lambda_i & \text{if } i \in I^0(x), \text{ where } \lambda_i \in [0,1]. \end{cases}$$

||

Note that for all $x, p \in \Re^n$, that $\|J_2(x)p\| \leq \|A\|\,\|p\|$, for all possible choices of $J_2(x) \in \partial(Ax-b)_+$. Commonly, $J_1(x)$ is assumed to be Lipschitz continuous of rank γ on some open convex set D. That is $\|J_1(x) - J_1(y)\| \leq \gamma\|x-y\|$ for all $x, y \in D$. The following proposition follows from real analysis and is found in most elementary material on continuous optimization, including Dennis and Schnabel[1983].

Proposition 3.3.3.

If J_1 is order 1 Lipschitz continuous of rank γ on an open set D, then

$$\|F_1(x+p) - F_1(x) - J_1(x)p\| \leq \frac{\gamma}{2}\|p\|^2,$$

where x and $x+p \in D$.

Proof. Since J_1 is Lipschitz continuous, it is differentiable a.e. on D, and the line integral parameterized along x to $x+p$ satisfies:

$$\begin{aligned} \|F_1(x+p) - F_1(x) - J_1(x)p\| &\leq \int_0^1 \|J_1(x+tp) - J_1(x)\|\,\|p\|\,dt \\ &\leq \int_0^1 \gamma\|tp\|\,\|p\|\,dt \\ &= \gamma\|p\|^2 \int_0^1 t\,dt \\ &= \frac{\gamma}{2}\|p\|^2. \end{aligned}$$

||

Observe that $J_2(x)$, a generalized Jacobian of $(Ax-b)_+$, need **not** satisfy a Lipschitz condition.

Proposition 3.3.4. *Existence of an Identification Neighborhood.*

Let x_* be given and let

$$I^0(x_*) := \{i \,|\, \langle a_i, x_* \rangle = b_i\},$$

$$I^+(x_*) := \{i \,|\, \langle a_i, x_* \rangle > b_i\},$$

$$I^-(x_*) := \{i \,|\, \langle a_i, x_* \rangle < b_i\}.$$

Then there exists a neighborhood $N(x_*, r)$ of x_*, such that for $x \in N(x_*, r)$:

$$\begin{aligned} I^0(x) &\subset I^0(x_*), \\ I^+(x_*) &= I^+(x) \backslash I^0(x_*), \\ I^-(x_*) &= I^-(x) \backslash I^0(x_*). \end{aligned} \tag{3.3.3}$$

Proof. Removing the constraints, $i \in I^0(x_*)$, allows forming a neighborhood, since there are then all strict inequalities. Define hyperplanes for i, $H_i := \{x | \langle a_i, x\rangle - b = 0\}$, and let $r > 0$ be chosen so that,

$$r < \min\{\text{ dist } (x_*, H_i) | \, i \in I^+(x_*) \cup I^-(x_*)\}.$$

If $x \in N(x_*, r)$, then

$$\begin{aligned} I^0(x) &\subset I^0(x_*), \\ I^+(x_*) &= I^+(x) \backslash I^0(x_*), \\ I^-(x_*) &= I^-(x) \backslash I^0(x_*). \end{aligned}$$

||

Proposition 3.3.5. *Perfect Approximating Property*

Let x_* be given and let $N(x_*, r)$ be a neighborhood which satisfies (3.3.3) in Proposition 3.3.4. Let $x \in N(x_*, r)$ and let $p := x_* - x$, then

1. for any $J_2(x) \in \partial(Ax - b)_+$,

$$(A(x+p) - b)_+ - (Ax - b)_+ - J_2(x)p = 0;$$

2. for any $J_2(x_*) \in \partial(Ax_* - b)_+$,

$$\|(Ax - b)_+ - (Ax_* - b)_+ - J_2(x_*)(-p)\| \le \|A_{I^0(x_*)}\| \, \|p\|;$$

3. for each particular $x \in N(x_*, r)$, there is at least one $J_2(x_*) \in \partial(Ax_* - b)_+$, such that

$$(Ax - b)_+ - (Ax_* - b)_+ - J_2(x_*)(-p) = 0.$$

Proof. For 1, note that $x + p = x_*$. Consider any i:

1.1. If $i \notin I^0(x_*)$. This implies $i \notin I^0(x)$. If $i \in I^-(x) \cup I^+(x)$, then $(J_2(x))_i$ is unique. In either case, then

$$(\langle x + p, a_i\rangle - b_i)_+ - (\langle x, a_i\rangle - b_i)_+ - (J_2(x))_i p = 0.$$

1.2. If $i \in I^0(x_*)$. If $i \in I^-(x) \cup I^+(x)$, then $(J_2(x))_i$ is unique. If $i \in I^0(x)$, then $\langle p, a_i\rangle = 0$, so for any $(J_2(x))_i = \lambda a_i$, where $\lambda \in [0, 1]$, $\langle p, \lambda a_i\rangle = 0$. In any case, then

$$(\langle x + p, a_i\rangle - b_i)_+ - (\langle x, a_i\rangle - b_i)_+ - (J_2(x))_i p = 0.$$

Relation (2) says that all the error of the first order approximation is in the components of the indices $i \in I^0(x_*)$, for this neighborhood.

For 3, Let $x \in N(x_*, r)$, and let

$$I^+ := I^+(x) \cap I^0(x_*),$$

$$I^- := I^-(x) \cap I^0(x_*).$$

The only choices are for those $i \in I^0(x_*)$. Choose

$$(J_2(x_*))_i = \begin{cases} a_i & \text{for } i \in I^+, \\ 0 & \text{for } i \in I^-, \\ \lambda a_i & \text{for any } \lambda \in [0,1],\ i \in I^0(x_*) \cap I^0(x). \end{cases}$$

Then $J_2(x_*) \in \partial(Ax_* - b)_+$ and

$$(Ax - b)_+ - (A(x_*) - b)_+ - J_2(x_*)(-p) = 0.$$

||

From Proposition 3.3.5, one can guarantee a type of perfect first order local approximation. For any x in this Proposition 3.3.4 neighborhood of x_*, any generalized Jacobian selection of $\partial(Ax - b)_+$ can be used in the first order Taylor series to perfectly first order model $(Ax^* - b)_+$ from the point x. On the other hand, the reverse may not hold. There may not be any generalized Jacobian selection of $\partial(Ax_* - b)_+$ for which the Taylor series based on this generalized Jacobian perfectly first order models $(Ax - b)_+$, for all x from this point x_*. But for each x, there is at least one generalized Jacobian in $\partial(Ax_* - b)_+$, for which the Taylor series using this generalized Jacobian models $(Ax - b)_+$ perfectly. The local approximating property of Proposition 3.3.5., the first result, is the critical property to derive properties of local convergence to a stationary point, such as local convergence rates.

This neighborhood is called the *identification* neighborhood of a point x_*. In this neighborhood, a point x has the same strict inequalities as for x_*, with the exception of the indices for $i \in I^0(x_*)$, where a containment property holds, $I^0(x) \subset I^0(x_*)$. Other research has addressed this neigborhood. Han[1980], Burke[1983], and Burke and Moré[1988] use this property in analyzing algorithms to solve systems of linear inequalities. They state an equivalent form of the property for the linear case, stating that the directions stay in the null space of the space spanned by the coefficients of the active constraints of the stationary point x_*. They did not develop this property in the setting of generalized differential constructs. In section 5, we address the class of C^2 convex functions and develop the local approximation results for this class of functions, applying the approximation result in a unified way for common algorithms. Robinson[1990] refers to the approximating property (3.3.5) as a point-based linearization approximating property, from which he develops local convergence properties of a Newton-like method to solve for zeroes of nonsmooth functions. Robinson does not show the details of how to form the point-based approximation for specific functions, rather he applies this approximating property to analyze general convergence properties.

Denote any generalized Jacobian of F at x, denoted $J(x)$, to be

$$J(x) = \begin{pmatrix} J_1(x) \\ J_2(x) \end{pmatrix}, \text{ where } J_2(x) \in \partial(Ax - b)_+.$$

f_2, being a differentiable convex function, is continuously differentiable.

Proposition 3.3.6.

∇f_2 is Lipschitz continuous of rank $\|A\|^2$ of order 1 over all $\Re^n$.

Proof. For all x, y,

$$\begin{aligned}\|\nabla f_2(x) - \nabla f_2(y)\| &= \|A^T(Ax-b)_+ - A^T(Ay-b)_+\| \\ &\leq \|A\| \, \|(Ax-b)_+ - (Ay-b)_+\| \\ &\leq \|A\|^2 \|x-y\|,\end{aligned}$$

where the last inequality comes from Proposition 3.3.1. $\|$

Therefore, one can also derive a Clarke generalized Jacobian set of ∇f_2, or a *generalized Hessian set*, denoted $\partial \nabla f_2(x)$. Let Ω be the set of points at which $\nabla f_2(x)$ is not differentiable. Then $\partial \nabla f_2(x)$ is the set all $n \times n$ matrices, $H_2(x)$, where $H_2(x)$ is a convex combination of a finite number of limits, H_2^k, where $k = 1, \ldots, K$, some K. Each limit H_2^k is obtained as the limit of a sequence of the form $\nabla^2 f_2(y_l)$ for some sequence $y_l \rightarrow x$ where $y_l \notin \Omega$. That is,

$$H_2^k := \lim_{y_l \rightarrow x} \nabla^2 f_2(y_l) \text{ where } y_l \notin \Omega.$$

Lemma 1.

If x_* is given, then there exists a neigborhood $N(x_*, r)$ of x_*, such that for $x \in N(x_*, r)$ and any $J_2(x) \in \partial(Ax-b)_+$ and $J_2(x_*) \in \partial(Ax_* - b)_+$,

$$(J_2(x) - J_2(x_*))^T (Ax_* - b)_+ = 0.$$

Proof. Let $I^+(x_*) := \{i | \langle a_i, x_* \rangle - b_i > 0\}$. If $I^+(x_*) = \emptyset$, then the result follows, so consider $I^+(x_*) \neq \emptyset$. Let $H_i := \{x | \langle a_i, x \rangle - b_i = 0\}$ be the hyperplane defined by a_i, b_i. Let $r > 0$ be chosen so that,

$$r < \min\{ \text{ dist } (x_*, H_i) | \, i \in I^+(x_*)\}.$$

If $x \in N(x_*, r)$, then

$$I^+(x_*) \subset I^+(x) = \{i | \langle a_i, x \rangle - b_i > 0\},$$

and for indices $i \in I^+(x_*)$, $(\langle a_i, x \rangle - b_i)_+$ is differentiable and

$$\partial(\langle a_i, x \rangle - b_i)_+ = a_i.$$

Then

$$(J_2(x) - J_2(x_*)^T (Ax_* - b)_+ = \sum_{i \in I^+(x_*)} (a_i - a_i)(Ax_* - b)_i = 0,$$

since for $i \notin I^+(x_*)$, $(\langle a_i, x_* \rangle - b_i)_+ = 0$. $\|$

Lemma 1 gives a neighborhood which is a superset of the identification neighborhood about x_*, since only the indices in $I^+(x_*)$ are used to define the neighborhood. This lemma is used in the proof of Theorem 1 in Section 4.

Proposition 3.3.7.

$\partial\nabla f_2(x) = \{J_2^T(x)J_2(x) \mid J_2(x) \in \partial(Ax-b)_+\}$.

Proof. Using the definition of the generalized Hessian, one can compute H_k obtained as the limit of a sequence of the form $\nabla^2 f_2(y_l)$, where $y_l \to x$ and $y_l \notin \Omega$ giving:

$$\begin{aligned} H_k &:= \lim_{y_l \to x} \nabla^2 f_2(y_l) \text{ where } y_l \notin \Omega \\ &= \lim_{y_l \to x} \mathcal{J}(\mathcal{J}_2^T(y_l))(Ay_l - b)_+ + \mathcal{J}_2^T(y_l)\mathcal{J}_2(y_l) \text{ where } y_l \notin \Omega \\ &= \lim_{y_l \to x} \mathcal{J}_2^T(y_l)\mathcal{J}_2(y_l) \text{ where } y_l \notin \Omega, \end{aligned}$$

since, for $i \in I^0(x) \cup I^-(x)$, $\lim_{y_l \to x}(\langle a_i, y_l\rangle - b)_+ = 0$, and for $i \in I^+(x)$, $(J_2(x))_i = (\mathcal{J}_2(x))_i = a_i$, a constant, which implies

$$\lim_{y_l \to x} \mathcal{J}(\mathcal{J}_2^T(y_l))(Ay_l - b)_+ = 0.$$

Therefore, a limit H_k is expressed: $H_k = (E_kA)^T E_kA$, for some E_k, which is an $m2 \times m2$ diagonal matrix with elements e_i^k on the diagonal, and

$$e_i^k = \begin{cases} 0 & \text{if } i \in I^-(x), \\ 1 & \text{if } i \in I^+(x), \\ 0 \text{ or } 1 & \text{if } i \in I^0(x). \end{cases} \tag{3.3.4}$$

Note that $E_k^2 = E_k$. Then

$$(E_kA)^T E_kA = (E_k^2 A)^T A = (E_kA)^T A.$$

First show that $\partial\nabla f_2(x) \subset \{J_2^T(x)J_2(x) | J_2(x) \in \partial(Ax-b)_+\}$. For λ_k satisfying $\sum_{k=1}^K \lambda_k = 1$ and $\lambda_k \in [0,1]$, then $H_2(x) \in \partial\nabla f_2(x)$ means:

$$\begin{aligned} H_2(x) &= \sum_{k=1}^K \lambda_k H_k \\ &= \sum_{k=1}^K \lambda_k (E_kA)^T E_kA \\ &= (\sum_{k=1}^K \lambda_k E_kA)^T A \\ &= ((\sum_{k=1}^K \lambda_k E_k)^{\frac{1}{2}}(\sum_{k=1}^K \lambda_k E_k)^{\frac{1}{2}} A)^T A \\ &= ((\sum_{k=1}^K \lambda_k E_k)^{\frac{1}{2}} A)^T (\sum_{k=1}^K \lambda_k E_k)^{\frac{1}{2}} A, \end{aligned}$$

since

$$\sum_{k=1}^K \lambda_k E_k = (\sum_{k=1}^K \lambda_k E_k)^{\frac{1}{2}}(\sum_{k=1}^K \lambda_k E_k)^{\frac{1}{2}},$$

because $\lambda_k \in [0,1]$, and E_k is diagonal with the property (3.3.4). Since $\sum_{k=1}^K \lambda_k = 1$ and $\lambda_k \in [0,1]$, then

$$(\sum_{k=1}^K \lambda_k E_k)^{\frac{1}{2}} A \in \partial(Ax-b)_+.$$

Now, to show containment in the other direction, express $J_2(x) = (\sum_{k=1}^{K} \lambda_k E_k)^{\frac{1}{2}} A$, solving for E_k which are diagonal and satisfy (3.3.4) and λ_k satisfying $\sum_{k=1}^{K} \lambda_k = 1$ and $\lambda_k \in [0,1]$. This is easily solved for indices $i \in I^+(x) \cup I^-(x)$, since then $(J_2(x))_i$ is unique (and e_i^k is unique) and any choice of $\lambda_k \in [0,1]$ where $\sum_{k=1}^{K} \lambda_k = 1$ works. For $i \in I^0(x)$, one must solve simultaneously for some finite K, and $e_i^k = 0$ or 1, and $\lambda_k \in [0,1]$ where $\sum_{k=1}^{K} \lambda_k = 1$:

$$(J_2(x))_i = ((\sum_{k=1}^{K} \lambda_k E_k)^{\frac{1}{2}} A)_i. \tag{3.3.5}$$

Since for $\lambda_k \in [0,1]$ where $\sum_{k=1}^{K} \lambda_k = 1$ implies $((\sum_{k=1}^{K} \lambda_k E_k)^{\frac{1}{2}})_i \in [0,1]$, and since the positive square root function is a bijection on $[0,1]$, then this system is underdetermined, implying (3.3.5) is solvable. Thus, one would reverse the arguments in the other direction, given any generalized Jacobian $J_2(x) \in \partial(Ax-b)_+$, then some $H_2(x) \in \partial\nabla f_2(x)$ exists su h that $H_2(x) = J_2^T(x)J_2(x)$. ||

Proposition 3.3.7 gives one convenient way to choose a generalized Hessian of $f_2(x)$ using any selection of $J_2(x) \in \partial(Ax-b)_+$. Note that $J_2^T(x)J_2(x)$ is symmetric, positive semi-definite. Also, for any $x, p \in \Re^n$ and any choice of $J_2(x) \in \partial(Ax-b)_+$, it follows that $|\langle p, J_2^T(x)J_2(x)p\rangle| \le \|A\|^2\|p\|^2$.

Proposition 3.3.8.

Let x_* be given and let $N(x_*, r)$ be a Proposition 3.3.4 identification neighborhood. Let $x \in N(x_*, r)$ and let $p := x_* - x$. Then

1. for any $J_2(x) \in \partial(Ax-b)_+$,

$$f_2(x+p) - f_2(x) - \langle p, \nabla f_2(x)\rangle - \frac{1}{2}\langle p, J_2^T(x)J_2(x)p\rangle = 0;$$

2. for any $J_2(x_*) \in \partial(Ax_* - b)_+$,

$$|f_2(x) - f_2(x_*) - \langle(-p), \nabla f_2(x_*)\rangle - \frac{1}{2}\langle p, J_2^T(x_*)J_2(x_*)p\rangle| \le \frac{1}{2}|\langle A_{I^0(x_*)}p, (A_{I^0(x_*)} + 2A_{I^+(x_*)})p\rangle|;$$

3. for each particular $x \in N(x_*, r)$ there is at least one $J_2(x_*) \in \partial(Ax_* - b)_+$, such that

$$f_2(x) - f_2(x_*) - \langle(-p), \nabla f_2(x_*)\rangle - \frac{1}{2}\langle p, J_2^T(x_*)J_2(x_*)p\rangle = 0.$$

Proof. For 1, Proposition 3.3.5 means that $F_2(x+p) = F_2(x) + J_2(x)p$. Therefore, taking inner products:

$$\langle F_2(x+p), F_2(x+p)\rangle = \langle F_2(x) + J_2(x)p, F_2(x) + J_2(x)p\rangle,$$
$$f_2(x+p) = f_2(x) + \langle J_2(x)p, F_2(x)\rangle + \frac{1}{2}\langle p, J_2^T(x)J_2(x)p\rangle.$$

For 2, this follows from the property that all the error is in the components having indices $i \in I^0(x_*)$ and is all contained in the term: $\frac{1}{2}\langle p, J_2^T(x_*)J_2(x_*)p\rangle$.

$$\frac{1}{2}\langle p, J_2^T(x_*)J_2(x_*)p\rangle =$$
$$\frac{1}{2}\left(\langle J_{I^0(x_*)}^2 p, J_{I^0(x_*)}^2 p\rangle + \langle A_{I^+(x_*)}p, A_{I^+(x_*)}p\rangle + 2\langle J_{I^0(x_*)}^2 p, A_{I^+(x_*)}p\rangle\right).$$

Therefore, since for $i \in I^0(x_*)$, $(J_2(x_*))_i = \lambda_i a_i$, for some $\lambda_i \in [0, 1]$, then

$$|f_2(x) - f_2(x_*) - \langle(-p), \nabla f_2(x_*)\rangle - \frac{1}{2}\langle p, J_2^T(x_*)J_2(x_*)p\rangle| \le \frac{1}{2}|\langle A_{I^0(x_*)}p, (A_{I^0(x_*)} + 2A_{I^+(x_*)})p\rangle|.$$

For 3, choose $J_2(x_*)$ as in Proposition 3.3.5. Then the result follows, just like in the first result of this proposition. $\|$

Denote $\Gamma_i(x)$ to be the Hessian matrix of $(F_1(x))_i$, and denote $g(x)$ as the gradient of $f(x)$. Denote $g_1(x)$, $g_2(x) = A^T(Ax - b)^+$, the gradients of $f_1(x)$ and $f_2(x)$, respectively. A choice $H(x) \in \partial\nabla f(x)$, a generalized Hessian of $f(x)$, is determined by the choice of generalized Jacobian of $(Ax - b)_+$. Then

$$\begin{aligned} g(x) &= J(x)^T F(x) = J_1^T F_1(x) + A^T(Ax - b)^+ = g_1(x) + g_2(x), \\ H(x) &= J(x)^T J(x) + Q(x) = J_1(x)^T J_1(x) + J_2(x)^T J_2(x) + Q(x), \\ \text{where } Q(x) &:= \sum_i^{m1} (F_1(x))_i \Gamma_i(x). \end{aligned} \tag{3.3.6}$$

4 Generalization of Algorithms for Solving a System of Linear Inequalities

In this section, we develop algorithms to solve problem (3.1.2). The specific algorithms are adaptations of the *Gauss-Newton and Levenberg Marquardt* methods, based on using a selection of the generalized Jacobian at each iterate. In section 4.1, the Gauss-Newton like algorithm is defined and local convergence properties are analyzed, and the conditions for local q-quadratic convergence are verified. The Levenberg-Marquardt type algorithm is developed and the conditions for iocal convergence are stated. In section 4.2, global convergence methods are discussed. The Armijo line search algorithm is analyzed. The global trust region method is adapted to solve (3.1.2) using a generalized Hessian model function, and the conditions for convergence are verified.

4.1 Local Convergence Results

The properties of the algorithm along with the properties of the specific function being minimized guarantee local convergence in some neighborhood of the solution or stationary point. We show local convergence properties for the undamped search directions, the solutions d to $Bd = -g(x)$, where $g(x) = \nabla f(x)$ and B is the matrix determined by the specific algorithm. See section 1.5.

4.1.1. Adaptation of Gauss-Newton Directions: Iterates x_k are generated by

$$x_{k+1} = x_k + d_k$$

where d_k, the search direction, solves

$$J_k^T J_k d_k = -g(x_k) = -J_k^T F_k,$$

where J_k is a generalized Jacobian of F at x_k. Local convergence of the Gauss-Newton method in some neighborhood of the solution to (3.1.2) is assured under the conditions of Theorem 1. Compare to Theorem 10.2.1 of Dennis and Schnabel [1983] for the smooth case.

Theorem 1. *Local Convergence of Gauss-Newton for Linear Inequalities*

Let f and F be defined as in section 3.1, and assume the following:

1. $f_1(x)$ is continuously differentiable in an open subset, $D \subset \Re^n$ and $J_1(x)$ is Lipschitz continuous on D of rank γ and order 1.
2. There exists $x_* \in D$, such that $g(x_*) = 0$.

   Assume there exists a neighborhood $N(x_*, r) \subset D$, such that the following are satisfied:
3. $\|J_1(x)\| \leq \alpha$, for all $x \in N(x_*, r)$.
4. There exists $\sigma \geq 0$, such that for all $x \in N(x_*, r)$,

$$\|(J_1(x) - J_1(x_*))^T F_1(x_*)\| \leq \sigma \|x - x_*\|. \tag{4.1.1.1}$$

5. λ, the smallest eigenvalue of $J^T(x_*)J(x_*)$, for any choice of generalized Jacobian, satisfies:

$$\lambda > \sigma.$$

Then, for any $c \in (1, \frac{\lambda}{\sigma})$, there exists $\epsilon > 0$, such that for all $x_0 \in N(x_*, \epsilon)$, the Gauss-Newton iterates

$$x_{k+1} := x_k - (J(x_k)^T J(x_k))^{-1} J(x_k)^T F(x_k),$$

are well-defined and converge to x_* and satisfy:

$$\|x_{k+1} - x_*\| \le \frac{c\sigma}{\lambda}\|x_k - x_*\| + \frac{c\alpha\gamma}{2\lambda}\|x_k - x_*\|^2, \tag{4.1.1.2}$$

$$\|x_{k+1} - x_*\| \le \frac{c\sigma + \lambda}{2\lambda}\|x_k - x_*\| < \|x_k - x_*\|. \tag{4.1.1.3}$$

Proof. There are two conditions different from those of Dennis and Schnabel. $J_2(x)$ may not be Lipschitz continuous on D, and $J_2^T(x)J_2(x)$ may not be unique. Lemma 1 and Proposition 3.3.5 show that there exists a neigborhood $N(x_*, r_*)$ of x_*, such that for $x \in N(x_*, r_*)$ and for any choice of generalized Jacobian of $F_2(x)$,

$$(J_2(x) - J_2(x_*))^T(Ax_* - b)_+ = 0,$$

$$F_2(x_*) - F_2(x) - J_2(x)(x_* - x) = 0.$$

This implies that $J_2^T(x)(F_2(x_*) - F_2(x) - J_2(x)(x_* - x)) = 0$. Choose $r := \min(r, r_*)$.

For the proof of the theorem, choose some fixed $c \in (1, \frac{\lambda}{\sigma})$. By an argument similar to Proposition 1.3.1, since $J^T(x_*)J(x_*)$ is invertible for any choice of generalized Jacobian, there is a neighborhood of invertibility, such that for any $x \in N(x_*, \epsilon_1)$ and for any choice of generalized Jacobian,

$$\|(J^T(x)J(x))^{-1}\| \le \frac{c}{\lambda},$$

since $c > 1$. Let

$$\epsilon := \min \left\{ r, \epsilon_1, \frac{\lambda - c\sigma}{c\alpha\gamma} \right\}.$$

Choose $x_0 \in N(x_*, \epsilon)$ and then

$$\|(J^T(x_0)J(x_0))^{-1}\| \le \frac{c}{\lambda}. \tag{4.1.1.4}$$

Using induction, consider the case for $k = 0$. Then at the first step, x_1 is well defined and

$$\begin{aligned}
x_1 - x_* &= x_0 - x_* - (J^T(x_0)J(x_0))^{-1}J^T(x_0)F(x_0) \\
&= (J^T(x_0)J(x_0))^{-1}[J^T(x_0)F(x_0) + J^T(x_0)J(x_0)(x_* - x_0)] \\
&= (J^T(x_0)J(x_0))^{-1}[J^T(x_0)F(x_*) - J^T(x_0)(F(x_*) - F(x_0) - J(x_0)(x_* - x_0))] \qquad (4.1.1.5) \\
&= (J^T(x_0)J(x_0))^{-1}[J^T(x_0)F(x_*) - J_1^T(x_0)(F_1(x_*) - F_1(x_0) - J_1(x_0)(x_* - x_0)) \\
&\quad + J_2^T(x_0)(F_2(x_*) - F_2(x_0) - J_2(x_0)(x_* - x_0))].
\end{aligned}$$

By Proposition 3.3.3, given the Lipschitz condition on J_1,

$$\|F_1(x_*) - F_1(x_0) - J_1(x_0)(x_* - x_0)\| \le \frac{\gamma}{2}\|x_0 - x_*\|^2. \tag{4.1.1.6}$$

Since $J^T(x_*)F(x_*) = 0$, and by Lemma 1, Proposition 3.3.5, and condition 4, then

$$\|J_1^T(x_0)F_1(x_*) - J_2^T(x_0)(F_2(x_*) - F_2(x_0) - J_2(x_0)(x_* - x_0))\| \le \sigma\|x_* - x_0\|. \tag{4.1.1.7}$$

By condition 3, that $\|J_1(x_0)\| \le \alpha$ and above, it follows that

$$\begin{aligned}\|x_1 - x_*\| &\le \|(J^T(x_0)J(x_0))^{-1}\| \left[\|J_1^T(x_0)F_1(x_*) - J_2^T(x_0)(F_2(x_*) - F_2(x_0) - J_2(x_0)(x_* - x_0))\|\right.\\ &\quad \left. + \|J_1(x_0)\|\, \|F_1(x_*) - F_1(x_0) - J_1(x_0)(x_* - x_0)\|\right]\\ &\le \frac{c}{\lambda}\left[\sigma\|x_0 - x_*\| + \frac{\alpha\gamma}{2}\|x_0 - x_*\|^2\right],\end{aligned}$$

which proves (4.1.1.2) for $k = 0$. From above and the choice of ϵ,

$$\begin{aligned}\|x_1 - x_*\| &\le \|x_0 - x_*\| \left[\frac{c\sigma}{\lambda} + \frac{c\alpha\gamma}{2\lambda}\|x_0 - x_*\|\right]\\ &\le \|x_0 - x_*\| \left[\frac{c\sigma}{\lambda} + \frac{\lambda - c\sigma}{2\lambda}\right]\\ &= \frac{c\sigma + \lambda}{2\lambda}\|x_0 - x_*\|\\ &< \|x_0 - x_*\|,\end{aligned}$$

which proves (4.1.1.3) for $k = 0$. The induction step is proven in similar fashion, with substitution of $k + 1$ for "1" and k for "0". ||

If $F_1(x_*) = 0$, then σ can be chosen to be 0, and the convergence rate is q-quadratic in this neighborhood. σ gives a measure of the size of the residual of F_1 at x_*. If F is linear in some neighborhood of x_*, then σ can be chosen to be 0, since in this neighborhood, $J(x) - J(x_*) = 0$.

4.1.2 Adaptation of Levenberg-Marquardt Direction: Given a generalized Jacobian J_k of F at each iterate x_k, the search direction d_k solves:

$$(J_k^T J_k + \mu_k I)d_k = -g(x_k) = -J_k^T F_k,$$

where μ_k is a non-negative scalar which is chosen so that, among other properties, $(J_k^T J_k + \mu_k I)$ is positive definite. This method is well-defined even when J_k does not have full rank. There is a sequence of iterates $\{x_k\}$, and there is a sequence of non-negative scalars $\{\mu_k\}$. There are different techniques to choose μ_k so that at each iteration, $(J_k^T J_k + \mu_k I)$ is positive definite. One technique is the trust region method, which is discussed in section 4.2. Theorem 2 below, using the same conditions as Theorem 1, and the condition that the μ_k are bounded, gives local convergence properties for Levenberg-Marquardt directions. See for example, Dennis and Schnabel [1983], Theorem 10.2.6 for the smooth case.

Theorem 2. *Local Convergence of Levenberg Marquardt for Linear Inequalities*

Let the conditions 1-5 of Theorem 1 be satisfied. Let the sequence of non-negative scalars, $\{\mu_k\}$, be bounded by $M > 0$. If $\sigma < \lambda$, then for any $c \in (1, \frac{\lambda+M}{\sigma+M})$, there exists $\epsilon > 0$, such that for all $x_0 \in N(x_*, \epsilon)$, the Levenberg-Marquardt iterates:

$$x_{k+1} := x_k - (J_k^T J_k + \mu_k I)^{-1} J_k^T F_k,$$

are well defined and obey:

$$\|x_{k+1} - x_*\| \leq \frac{c(\sigma + M)}{\lambda + M}\|x_k - x_*\| + \frac{c\alpha\gamma}{2(\lambda + M)}\|x_k - x_*\|^2,$$
$$\|x_{k+1} - x_*\| \leq \frac{c(\sigma + M) + \lambda + M}{2(\lambda + M)}\|x_k - x_*\| < \|x_k - x_*\|.$$

Proof. Choose a neighborhood $N(x_*, r)$ about x_* as in Theorem 1. Since $\sigma < \lambda$, choose

$$c \in \left(1, \frac{\lambda + M}{\sigma + M}\right).$$

Then $c > 1$. Choose ϵ_1 such that $J_0^T J_0 + \mu_0 I$ is invertible and satisfies for $x_0 \in N(x_*, \epsilon_1)$:

$$\|(J_0^T J_0 + \mu_0 I)^{-1}\| \leq \frac{c}{\lambda + M}.$$

Then let

$$\epsilon := \min\left\{r, \epsilon_1, \frac{\lambda + M - c(\sigma + M)}{c\alpha\gamma}\right\}.$$

The proof then follows the same pattern as Theorem 1. $\|$

4.2 Global Convergence Results

There are two basic approaches to assure that iterates, starting from any initial location, converge to a stationary point. These are the inexact line search methods and the trust region methods. In section 1.5, the Armijo line search properties and the global trust region method are described for the smooth case. It is usually appropriate to apply the line search method in combination with steepest descent and quasi-Newton search direction methods, since these methods produce descent directions, directions d such that $\langle d, g(x)\rangle < 0$. Global trust region methods derive the step (scaled direction). If the model function is the second order Taylor series approximation, then one might consider the direction to be a Newton-like direction, in that it is derived using Hessian information. If the model function uses $J^T(x)J(x)$ to approximate the Hessian, then direction is related to a Gauss-Newton direction. If the model function uses $J^T(x)J(x) + \mu I$ to approximate the Hessian, then the direction is related to the Levenberg-Marquardt direction.

4.2.1 Inexact Line Search: We analyze properties of the function f, given by (3.1.2), necessary for the Armijo line search to assure global convergence to a stationary point. See the Armijo line search algorithm, (1.5.2). Notice that the function f, given by (3.1.2), is bounded below by 0. Further, ∇f_2 is uniformly Lipschitz of rank $\|A\|^2$ on all $\Re^n$ A well known result, due to Wolfe [1969,1971], shows the conditions for the global line search to converge to a stationary point.

Theorem 3. *Wolfe's Global Line Search Theorem. Let ∇f be order 1 Lipschitz continuous on $\{x| f(x) < f(x_0)\}$ of rank γ. Suppose an algorithm produces Armijo stepsizes, λ_k, with descent directions d_k, satisfying for all k:*

$$\langle \nabla f(x_k), \lambda_k d_k\rangle < 0 \quad \text{or} \quad d_k = 0.$$

If f is bounded below and the angle between d_k and $-\nabla f(x_k)$ is uniformly bounded away from orthogonality then, either:

$$\nabla f(x_k) = 0, \text{ some } k \quad \text{or} \quad \lim_{k\to\infty} \nabla f(x_k) = 0.$$

The function f, given by (3.1.2), satisfies the conditions of Wolfe's Theorem, if ∇f_1 is Lipschitz continuous on $\{x| f(x) < f(x_0)\}$ of some given rank, say γ_1. Therefore, no adjustment is needed in the global line search method to account for the generalized Jacobian of f_2. If a search direction method is chosen which assures descent directions, and if the method assures that the directions and the gradients are uniformly bounded away from orthogonality, then the Armijo line search assures global convergence to a stationary point. Such direction methods could be Levenberg-Marquardt or a positive definite quasi-Newton method.

4.2.2 Global Trust Region Method: See (1.5.3) and (1.5.4). To analyze the global trust region method, one considers the model function and how accurately it approximates the function, f. However, higher accuracy implies more difficult computations to form and solve subproblem (1.5.3). We consider two model functions, both using generalized Jacobians for forming the model. Suppose $H(x)$ is a generalized Hessian of $f(x)$, and $J(x)$ is a generalized Jacobian of $f(x)$. Let

$$M(s) := f(x) + \langle g(x), s\rangle + \frac{1}{2}\langle s, H(x)s\rangle \quad \text{be the Hessian (or Newton) model,}$$
$$L(s) := f(x) + \langle g(x), s\rangle + \frac{1}{2}\langle s, J^T(x)J(x)s\rangle \quad \text{be the Gauss-Newton model.}$$

Consider the generalized Hessian model first. The method finds a specific step s to satisfy agreement between the model function and the function f. An initial $\delta_1 > 0$ and x_1 is chosen. Using the current iterate function value, gradient, and the generalized Hessian at x_k, form the model function $M_k(s)$. The subproblem below is solved for s_k:

$$\min_s \quad M_k(s) \quad \text{subject to } \|s\| \le \delta_k. \tag{4.2.2.1}$$

For theoretical convergence results, assume that the subproblem is solved exactly. s_k gives the step which minimizes the Hessian model (4.2.2.1) in a bounded ball about x_k. How does the reduction in the f compare with the reduction in the model function? Calculate the ratio:

$$r_k := \frac{f(x_k) - f(x_k + s_k)}{M_k(0) - M_k(s)} = \frac{f(x_k) - f(x_k + s_k)}{f(x_k) - M_k(s_k)}. \tag{4.2.2.2}$$

Define the step parameters and the trust region size parameter, the same as in section 1.5, denoting the iterate subscript k. Let

$$\Delta f_k := f(x_k) - f(x_k + s_k),$$
$$\Delta M_k := M_k(0) - M_k(s_k).$$

1) If $r_k < .25$, then set $\delta_{k+1} := \|s_k\|/4$.
   If $r_k > .75$ and $\|s_k\| = \delta_k$, then set $\delta_{k+1} := 2\delta_k$,
   otherwise, set $\delta_{k+1} := \delta_k$.
2) If $r_k \le 0$, then set $x_{k+1} := x_k$; else, set next $x_{k+1} := x_k + s_k$.

Steps which reduce the objective function are called successful steps. The others are called null steps. If $r_k < .25$, then the step is called a region-reducing step. One may consider choosing a different generalized Hessian, if there is a choice, when taking a null step or a region reducing step. If there is an infinite sequence of iterates, then either the region reducing steps have an infinite subsequence or the steps, where $r \ge .25$, have an infinite subsequence or both. Suppose there is an infinite number of iterates. If $\delta_k \to 0$, then an infinite number of iterates must satisfy, $r_k < .25$, in order that $\delta_k \to 0$, meaning that there must be an infinite subsequence of region reducing steps. If there is an infinite number of iterates and $\inf_k \delta_k > 0$, then there must be an infinite subsequence of iterates, k, where $r_k \ge .25$.

A first order condition for x_* to solve (3.1.2) is that $\nabla f(x_*) = 0$, or, equivalently, for all p, $\langle \nabla f(x_*), p\rangle \succeq 0$. A point x_* satisfies second order minimality conditions if for every generalized Hessian at x_* and every p, $\langle p, H(x_*)p\rangle \ge 0$. Define "little o" notation, $o(\bullet)$, which means that $a = o(h) \iff \frac{a}{h} \to 0$, as $h \to 0$. We use this notation to discuss order of convergence, such as linear or quadratic.

Proposition 4.2.2.1.

Let S be a bounded set. Let f_1 be C^2 on $\Re^n$, and $\nabla^2 f_1$ be Lipschitz continuous of rank γ on S. Let $x_*, \{x_k\}_{k=1}^{\infty} \in S$, for all k, and let $\|s_k\| \to 0$ and $x_k \to x_*$. Then there exists an infinite subsequence of iterates x_k, $k \in \mathcal{L}$ satisfying,

$$|f(x_k + s_k) - M_k(s_k)| \leq \frac{\gamma}{6}\|s_k\|^3.$$

Proof. $\mathcal{L}$ denotes the set of indices of the subsequence. Since f_1 is C^2, and its Hessian is Lipschitz of rank γ on S, one can approximate f_1 second order with the first three terms of the Taylor series. This property is easily verified. Given γ, the Lipschitz constant for $\nabla^2 f_1$ on S, let $x, x+p \in S$, then

$$|f_1(x+p) - f_1(x) - \langle \nabla f_1(x), p\rangle - \frac{1}{2}\langle p, \nabla^2 f_1(x)p\rangle| \leq$$
$$\int_0^1 \left(\int_0^1 \|\nabla^2 f_1(x+tp) - \nabla^2 f_1(x)\|\, \|tp\|\, dt\right) \|p\|\, dt \leq$$
$$\int_0^1 \left(\int_0^1 \gamma\|tp\|\,\|p\|\, dt\right)\|p\|\, dt \leq$$
$$\int_0^1 \frac{\gamma}{3} t\|p\|^3\, dt \leq \frac{\gamma}{6}\|p\|^3.$$

For the proof of the proposition, since $s_k \to 0$, let K_1 be sufficiently large such that for some $k \geq K_1$, Proposition 3.1 assures an identification neighborhood $N(x_k + s_k, r_k)$ about $x_k + s_k$, such that $x_k \in N(x_k + s_k, r_k)$. By hypothesis that $\nabla^2 f_1$ is Lipschitz continuous of rank γ on S and the above result, then

$$|f_1(x_k + s_k) - f_1(x_k) - \langle \nabla f_1(x_k), s_k\rangle - \frac{1}{2}\langle s_k, \nabla^2 f_1(x)s_k\rangle| \leq \frac{\gamma}{6}\|s_k\|^3.$$

By Proposition 3.3.8, for $x_k \in N(x_k+s_k, r_k)$ and for any selection of generalized Hessian $H_2(x) = J_2^T(x)J_2(x)$, then

$$f_2(x_k + s_k) - f_2(x_k) - \langle s_k, \nabla f_2(x_k)\rangle - \frac{1}{2}\langle s_k, J_2^T(x_k)J_2(x_k)s_k\rangle = 0.$$

This implies for this $k > K_1$, that the part of the model for f_2, denoted M_2^k, satisfies:

$$M_2^k(s_k) = f_2(x_k + s_k).$$

Therefore,

$$|f(x_k + s_k) - M_k(s_k)| \leq \frac{\gamma}{6}\|s_k\|^3.$$

Let this $k \in \mathcal{L}$.

Now, since $\{x_k\}_{k=1}^{\infty}$ is an infinite sequence, $\{x_k\}_{k=K_1+1}^{\infty}$ is an infinite subsequence. Since $\|s_k\| \to 0$, find K_2 sufficiently large such that for some $k \geq K_2 > K_1$ one can again form an identification neighborhood as above. Repeat this process, forming an infinite subsequence of indices $k \in \mathcal{L}$. $\|$

Proposition 4.2.2.1 means that $f(x_k + s_k) - M_k(s_k) = o(\|s_k\|^2)$ for $k \in \mathcal{L}$.

We state and prove the global convergence theorem, which is a generalization of Fletcher[1987].

Theorem 4. *Generalized Hessian Model Global Trust Region Convergence*

Let f be given as in section 3.1. For the trust region method applied to the generalized Hessian model, let $\{x_k\}_{k=1}^{\infty} \in S$, where S is bounded. If f_1 is C^2 on S, and its Hessian is Lipschitz of rank γ on S, then there exists an accumulation point x_* which satisfies first and second order optimality conditions.

Proof. If $\inf \delta_k = 0$, then by the discussion above, there must be an infinite subsequence of region reducing steps, and hence, steps where $r_k < .25$. Denote $\mathcal{L}_1 := \{k | r_k < .25\}$, where $\inf \delta_k = 0$. If $\inf \delta_k > 0$, there must be an infinite subsequence of iterates such that $r_k \geq .25$. Denote $\mathcal{L}_2 := \{k | r_k \geq .25\}$. At least one of the two cases occurs, and by S being bounded, the infinite subsequence has an accumulation point, which is denoted x_*. Consider the two cases:

1. $r_k < .25$, $\delta_k \to 0$ and, hence $\|s_k\| \to 0$, $k \in \mathcal{L}_1$,
2. $r_k \geq .25$ and $\inf \delta_k > 0$, $k \in \mathcal{L}_2$.

Case 1: Let g_* denote $\nabla f(x*)$. Argue by contradiction. Suppose there is a descent direction p at x_*, with $\|p\| = 1$ such that

$$\langle p, g_* \rangle = -\alpha < 0. \tag{4.2.2.3}$$

By Proposition 4.2.2.1, there exists an infinite subsequence of iterates x_k, $k \in \mathcal{L} \subset \mathcal{L}_1$, such that for $k \in L$

$$\begin{aligned} f(x_k + s_k) &= M_k(s_k) + o(\|s_k\|^2), \\ \triangle f_k &= \triangle M_k + o(\|s_k\|^2). \end{aligned} \tag{4.2.2.4}$$

Consider a step of length $\epsilon_k = \|s_k\|$ along p. Since for any choice of generalized Hessian, H_k is uniformly bounded on S, and by optimality of (4.2.2.1), it follows:

$$\begin{aligned} \triangle M_k \geq M_k(0) - M_k(\epsilon_k p) &= -\langle g_k, \epsilon_k p \rangle - \langle \epsilon_k p, H_k \epsilon_k p \rangle \\ &= -\langle g_k, \epsilon_k p \rangle + o(\epsilon_k) \\ &= \epsilon_k \alpha + o(\epsilon_k), \end{aligned}$$

by continuity of $\langle \bullet, \bullet \rangle$ and (4.2.2.3). Forming the quotient r_k, since $\epsilon_k \alpha + o(\epsilon_k) \leq \triangle M_k$, for $k \in \mathcal{L}$, then

$$\begin{aligned} r_k &:= \frac{\triangle f_k}{\triangle M_k} \\ &= \frac{\triangle M_k + o(\epsilon_k^2)}{\triangle M_k} \\ &= 1 + \frac{o(\epsilon_k^2)}{\triangle M_k}. \end{aligned}$$

Since $\epsilon_k \to 0$ and by (4.2.2.4), then $r_k \to 1$, which contradicts that $r_k < .25$. Thus, it must be that $g_* = 0$.

Again, arguing by contradiction, if not all generalized Hessians at x_* are positive semi-definite, then there must be at least one second order descent direction. We choose a specific normalized descent direction d, along with a particular generalized Hessian H_* at x_*, for which $\langle d, H_* d \rangle < 0$ is furthest from zero. Let d or its negative be a descent direction with $\|d\| = 1$, and the generalized Hessian H_* at x_* solve:

$$\begin{aligned} \min_{p, H(x_*)} \quad & \langle p, H(x_*) p \rangle \\ \text{subject to:} \quad & H(x_*) \in \partial \nabla f(x_*) \\ & \|p\| = 1. \end{aligned}$$

A solution pair (d, H_*) for the above problem exists, because of the compactness of the generalized Hessian set, $\partial\nabla f(x_*)$, and since we assume at least one such direction and generalized Hessian exists, by our argument by contradiction. Therefore, for the solution (d, H_*) of above, $\langle d, H_* d\rangle = -\alpha$ for some $\alpha > 0$. Then for all $H(x_*)$ and for all directions p, such that $\|p\| = 1$,

$$\langle p, H(x_*)p\rangle \geq \langle d, H_* d\rangle = -\alpha < 0. \tag{4.2.2.5}$$

Notice for $k \in \mathcal{L} \subset \mathcal{L}_1$, for any other generalized Hessian of f_2, denoted H_2', and its associated model function, denoted $M_k^{2'}$, at $x_k + s_k$, that

$$f_2(x_k + s_k) = M_k^2(s_k) = M_k^{2'}(s_k).$$

For $k \in \mathcal{L} \subset \mathcal{L}_1$, choose a step length of ϵ_k, for direction σd, where $\sigma = \pm 1$, so that $\langle g_k, \sigma d\rangle \leq 0$. Notice that σd also solves (4.2.2.5). Now, by the upper-semicontinuity of the generalized Hessian map, $H_k \to H(x_*) \in \partial\nabla f(x*)$. Then by optimality of (4.2.2.1), and by the upper semi-continuity of the generalized Hessian,

$$\begin{aligned}
\triangle M_k &\geq M_k(0) - M_k(\epsilon_k \sigma d) \\
&\geq -\frac{1}{2}\epsilon_k^2\langle d, H_k d\rangle \\
&= -\frac{1}{2}\epsilon_k^2\langle d, H(x_*)d\rangle + o(\epsilon_k^2) \\
&\geq \frac{1}{2}\epsilon_k^2\alpha + o(\epsilon_k^2).
\end{aligned}$$

Again, one can form the quotient ratio r_k, as above, and from (4.2.2.4), then $r_k \to 1$ which contradicts that $r_k < .25$. Thus, H_* is positive semidefinite, and any choice of generalized Hessian at x_* is also positive definite. That is, for any direction p, such that $\|p\| = 1$, it follows that, for any $\lambda \geq 0$

$$\begin{aligned}
&0 \leq \langle d, H_* d\rangle \leq \langle p, H(x_*)p\rangle \\
&0 \leq \langle \lambda d, H_* \lambda d\rangle \leq \langle \lambda p, H(x_*)\lambda p\rangle.
\end{aligned}$$

Case 2: In this case, there is a subsequence $k \in \mathcal{L}_2$, where $f_1 - f_* = \sum_{k=1}^{\infty}(f_k - f_{k+1}) \geq \sum_{k\in\mathcal{L}_2}\triangle f_k \geq .25\sum_{k\in\mathcal{L}_2}\triangle M_k$, since $r_k \geq .25$ for $k \in \mathcal{L}_2$. Since f is bounded on S and $x_* \in S$, then $f_1 - f_*$ is uniformly bounded, and it must be that $\triangle M_k \to 0$. Let $\bar\delta$ satisfy: $0 < \bar\delta < \inf \delta_k$. Let $\bar s$ and H_*, a particular generalized Hessian of f at x_*, solve:

$$\begin{aligned}
\min_{s, H(x_*)} \quad & \langle s, H(x_*)s\rangle \\
\text{subject to:} \quad & H(x_*) \in \partial\nabla f(x_*) \\
& \|s\| \leq \bar\delta.
\end{aligned} \tag{4.2.2.6}$$

(4.2.2.6) has a solution since the generalized Hessian set is compact. Define $M_*(s) := f_* + \langle s, g_*\rangle + \frac{1}{2}\langle s, H_* s\rangle$. Define $\bar x := x_* + \bar s$. For sufficiently large k, by choice of $\bar\delta$, and since x_* is a cluster point of $\{x_k | k \in \mathcal{L}_2\}$, then

$$\|\bar x - x_k\| \leq \|\bar s\| + \|x_k - x_*\| =\leq \bar\delta + \|x_k - x_*\| \leq \delta_k.$$

This implies $\bar{x} - x_k$ is feasible for the subproblem (4.2.2.1). Thus,

$$M_k(\bar{x} - x_k) \geq M_k(s_k) = f_k - \triangle M_k,$$

since s_k solves the subproblem (4.2.2.1). In the limit, $f_k \rightarrow f_*$, $g_k \rightarrow g_*$, $\triangle M_k \rightarrow 0$, and $\bar{x} - x_k \rightarrow \bar{s}$. Since H_* solves (4.2.2.6), then for any other $H(x_*)$, define

$$M_*'(\bar{s}) := f_* + \langle \bar{s}, g_* \rangle + \frac{1}{2}\langle \bar{s}, H(x_*)\bar{s} \rangle \geq M_*(\bar{s}),$$

and then

$$M_*'(\bar{s}) \geq M_*(\bar{s}) \geq f_* = M_*(0) = M_*'(0).$$

Thus, $s = 0$ also minimizes $M_*(s)$ on $\|s\| \leq \bar{\delta}$. Since the constraints are not active at $s = 0$, then first and second order optimality conditions must be satisfied for the subproblem (4.2.2.1). That is, $g_* = 0$, and H_* must be positive semi-definite. Therefore, on $\|s\| \leq \bar{\delta}$ and for any generalized Hessian $H(x_*)$ at x_*, then

$$0 \leq \langle \bar{s}, H_*, \bar{s} \rangle \leq \langle s, H(x_*)s \rangle.$$

This implies that $H(x_*)$ is also positive semi-definite. ||

If the lower level set of the first iterate is bounded and since the function values at the iterates decrease or stay the same, then he existence of the bounded set S required for the theorem is satisfied. The coercive properties of a function for which this is satisfied are discussed in section 4.3. We show in section 4.3, that this condition is satisfied for the specific function f determined by (2.5.1). If the sequence of iterates are finite, that is $\triangle M_k = 0$ for some k, then first and second order optimality conditions are satisfied, by an argument similar to the one in case 2 of Theorem 4.

We now consider the Levenberg-Marquardt method based on a trust region method. Let $x_k \in \Re^n$ and $J(x_k)$ be a generalized Jacobian of f at x_k. The generalized Jacobian model function for the non-linear least squares problem is:

$$M(x_k + s_k) := f(x_k) + \langle \nabla f(x_k), s_k \rangle + \frac{1}{2}\langle s_k, J^T(x_k)J(x_k)s_k \rangle.$$

Let

$$B_k := J(x_k)^T J(x_k).$$

In solving for the model trust region, one wishes to find the direction s_k of norm less than or equal to some $\delta > 0$, such that in this direction, there is agreement between the decrease in the model function and the decrease in f. Find s_k which solves:

$$\begin{aligned} \min_{s} \quad & f(x_k) + \langle \nabla f(x_k), s \rangle + \frac{1}{2}\langle s, J^T(x_k)J(x_k)s \rangle \\ \text{subject to: } & \langle s, s \rangle \leq \delta^2. \end{aligned} \tag{4.2.2.7}$$

The following proposition shows the conditions for optimality and uniqueness of a solution of (4.2.2.7). This proposition is found in most elementary material. See, for example, Fletcher[1987].

Proposition 4.2.2.2.

s_k solves (4.2.2.7) $\iff$ $\langle s_k, s_k\rangle \le \delta^2$ and there exists $\mu \ge 0$, such that

$$\begin{aligned}(B_k + \mu I)s_k &= -\nabla f(x_k),\\ \mu(\langle s_k, s_k\rangle - \delta^2) &= 0.\end{aligned} \tag{4.2.2.8}$$

Furthermore, if $B_k + \mu I$ is positive definite, then s_k is the unique minimizer.

Proof: Since B_k is positive semi-definite, (4.2.2.7) is a convex problem. (4.2.2.8) are the KKT conditions. The uniquenss of s_k is assured by the strict convexity of the Lagrangian of (4.2.2.7):

$$\langle \nabla f(x_k), s\rangle + \frac{1}{2}\langle s, (B_k + \mu I)s\rangle.$$

||

For the case, where μ is chosen so that $B_k + \mu I$ is positive definite, there is a unique s_k. Given a specific δ, there is an associated μ which solves (4.2.2.7). The larger the δ, then the smaller that μ is. Thus, the problem becomes: find the δ, as large as possible, while still assuring $B_k + \mu I$ is positive definite. The trust region algorithms do not solve for μ exactly. Instead, an approximation process is used, where δ is decreased or increased using the trust region parameters outlined for the generalized Hessian model. This approach compares the changes in the model function and the function f, and adjusts δ and s_k appropriately, the same as for the generalized Hessian trust region method.

By choosing $J_2(x_k)$ appropriately, one may avoid the need for $\mu_k > 0$ at some iterates. For the function f given by (3.1.2), choose a specific generalized Jacobian at x_k, denoted J_2^k where for indices $i \in I^0(x_k)$, choose $(J_2^k)_i = a_i$. If for a direction s, such that $\langle s, a_i\rangle \ne 0$ for at least one $i \in I^0(x_k)$, then for any other choice of generalized Jacobian $J_2(x_k)$,

$$\langle J_2^k s, J_2^k s\rangle > \langle J_2(x_k)s, J_2(x_k)s\rangle.$$

The model function based on the Hessian approximation $J^T(x_k)J(x_k)$ gives a positive semi-definite model. Thus, in some cases, this choice of generalized Jacobian which gives an increase in positive definiteness may be sufficient to assure:

$$\langle J_1(x)s, J_1(x)s\rangle + \langle J_2^k s, J_2^k s\rangle > 0,$$

compensating for lack of positive definiteness in the function f_1 at x_k. This strategy of choosing the most positive definite choice of $J^T(x_k)J(x_k)$ among generalized Jacobians of $(Ax_k - b)_+$ increases the possibility of avoiding $\mu_k > 0$ for some iterates.

If F_1 is highly non-linear near a stationary point, and if the residual is large, then using the Hessian approximation, $J^T J$, may not work. Varah[1990] uses statistical arguments to compare the relative sizes of the Hessian terms in the non-linear least square problem. He reports that even when the the norm of the term $Q(x)$ is relatively large with respect to the norm of $J^T J$, the algorithm using only $J^T J$ works well.

4.3 Recession Behavior of the Least Square Problem Formulation

Notice in the trust region algorithm, that the existence of an accumulation point of the sequence of iterates, x_k, depends on whether the iterates stay bounded. Since the least square objective function f is bounded below, we would like to characterize a property of f which assures, that as long as the function values at the iterates significantly decrease (by satisfying the line search conditions or the trust region conditions), that the sequence generates an accumulation point. One condition to avoid is that the function $f(x)$ recedes, that is does not increase, as $\|x\| \to \infty$. This property of a function is called *recession behavior*. As an example, a function such as $\frac{1}{x}$ recedes to 0 in the direction $d = 1$. Given $x_0 > 0$ and $\lambda > 0$,

$$\lim_{\lambda \to +\infty} \frac{1}{x_0 + \lambda d} = 0,$$

and $\frac{1}{x_0+\lambda d}$ is strictly decreasing as $\lambda \to +\infty$. Another property, contrasted to recession behavior, is called *coercive* behavior. Define a function as *coercive* if for all x,

$$\lim \quad f(x) = +\infty \qquad \text{as } \|x\| \to \infty.$$

Recall that solving for

$$\inf \quad \frac{1}{2}\langle (Ax-b)_+, (Ax-b)_+ \rangle \tag{4.3.1}$$

is equivalent to solving the quadratic programming problem:

$$\inf_{x,z} \quad \langle z, z \rangle, \tag{4.3.2}$$

$$\text{subject to:} \quad Ax - b \leq z,$$

and the optimality conditions for the unique z_* and solution(s) x which solve (4.3.2) are:

$$A^T z_* = 0,$$

$$Ax \leq b + z_*,$$

$$z_* \geq 0,$$

$$\langle z_*, Ax - b - z_* \rangle = 0.$$

Denote the solution set of (4.3.1) as $\Omega := \{x | \, Ax - b \leq z_* \text{ and } \langle z_*, Ax - b - z_* \rangle = 0\} \neq \emptyset$. Let $I := \{i | \, z_i^* > 0\}$ and $I' := \{i | \, z_i^* = 0\}$. Denote A_I, the matrix formed from the rows of A for the indices $i \in I$. Then

$$\Omega = \{x | \, A_I x = b_I + z_I^* \text{ and } A_{I'}x - b_{I'} \leq 0\}. \tag{4.3.3}$$

If d is some direction, such that for $x \in \Omega$ implies $x + \lambda d \in \Omega$, for all $\lambda \geq 0$, then:

$$A_I(x + \lambda d) = b_I + z_I^* \text{ and } A_{I'}(x + \lambda d) - b_{I'} \leq 0.$$

Then for all $\lambda \geq 0$:

$$A_I \lambda d = 0_I, \tag{4.3.4}$$

$$A_{I'}\lambda d \leq 0_{I'}. \tag{4.3.5}$$

(4.3.4) implies that $d \in \mathcal{N}(A_I)$, the null space of A_I. For (4.3.5), equivalently,

$$\alpha_{I'} \geq 0 \text{ implies } \langle d, A_{I'}^T \alpha \rangle \leq 0.$$

Let

$$K := \{A_{I'}^T \alpha \mid \alpha \geq 0_{I'}\}$$

be the cone generated by $A_{I'}^T$. Denote the polar of K,

$$K^* := \{d \mid \langle d, c \rangle \leq 0 \text{ for all } c \in K\}.$$

One can express the condition that d is a direction of recession of Ω, as follows:

$$d \in \mathcal{N}(A_I) \cap K^*. \tag{4.3.6}$$

Recall that $f_2(x) := \frac{1}{2}\langle (Ax-b)_+, (Ax-b)_+ \rangle$.

Proposition 4.3.1.

If $x \notin \Omega$ and $d \in \mathcal{N}(A_I) \cap K^*$, then there exists $\lambda_0 \geq 0$, such that for $\lambda \geq \lambda_0$, then $f_2(x+\lambda d) = f_2(x+\lambda_0 d)$ and $\langle \nabla f_2(x+\lambda d), d \rangle = 0$.

Proof. For $i \in I$, it follows that $A_I(x+\lambda d) = A_I x$, for any choice of λ. Let $r_{I'} := A_{I'} d \leq 0$. For $i \in I'$, such that $r_i < 0$, choose λ_0 to satisfy:

$$\langle x + \lambda_0 d, a_i \rangle - b_i \leq 0.$$

For $\lambda \geq \lambda_0$, then $f_2(x+\lambda d) = f_2(x+\lambda_0 d)$. The second part of the proposition, that for $\lambda \geq \lambda_0$, $\langle \nabla f_2(x + \lambda d), d \rangle = 0$ is implied by the property that $f_2(x+\lambda d) = f_2(x+\lambda_0 d)$. ||

Recall that $f_1(x) := \frac{1}{2}\langle F_1(x), F_1(x) \rangle$. Since f_1 is bounded below by 0, then $\liminf_y f_1(y) \geq 0$. There may be directions along which which f_1 recedes to some β, $\beta \geq 0$, never reaching it. Define a *recession direction* d of f_1 at some x, if

$$\liminf_{\lambda > 0} f_1(x+\lambda d) = \lim_{\lambda \to \infty} f_1(x+\lambda d) < \infty$$

The functions f_1 and f_2 from the problem formulated from the assignment problem in section 2 do not have a common direction of recession. The function $f_1(x) = \sum_j w_j^2(1-w_j)^2$ is a coercive function. f_2 is also coercive, since the polyhedral set defined by (2.3.1) is a polytope. Both f_1 and f_2 are bounded below. Therefore, if steps of sufficient size are taken, with sufficient reduction in the objective function to satisfy the Armijo conditions, and the angle between the step and gradient stays bounded away from orthogonal, and if there are an infinite number of iterates, then the sequence of iterates must have a cluster point. Similarly, for the trust region conditions to satisfy successful steps for an infinite number of iterates, there must be a cluster point in the sequence of successful iterates.

5 Properties of Least Square Problem for Solving a System of Convex Inequalites

One can extend the results for systems of linear inequalities to more general functions. This section defines the least square formulation for solving finite systems of inequalities and develops approximating properties of generalized differential constructs in modeling of convex C^2 functions. The results are similar to those for the case of linear inequalities developed in section 3. In section 5.1, the least square problem is defined for this case of convex C^2 inequalities, and the numerical stability of this problem is discussed. In section 5.2, the approximating properties are developed. The existence of an identification neighborhood is verified. The details for selecting generalized Jacobians and Hessians are addressed. Using the identification neighborhood property, several approximating results are developed using models based on generalized differential constructs. In the case of a finite system of inequalities of non-linear C^2 convex functions, the generalized Hessian model of the least square objective function does not perfectly approximate in the identification neighborhood of a designated point (See Proposition 3.3.8). However, the generalized Hessian model does provide second order approximation in an identification neighborhood of a designated point. Similar results are obtained for the generalized Jacobian model of the mapping F_+, as well as for the generalized Hessian model for ∇f. These approximating results are used in the development and analysis of the algorithms in section 6.

5.1 Statement of Problem

Suppose $F_i : R^n \mapsto R$, $i = 1, \ldots, m$ are convex functions, which are C^2. We wish to find x which solves:

$$F(x) \leq 0,$$

in a least square sense. Define

$$f(x) := \frac{1}{2}\langle F_+(x), F_+(x)\rangle.$$

Notice that f is a finite valued convex function, being the sum of convex functions (the pointwise maxima of the individual convex functions and the constant function zero). We wish to find x which solves:

$$\inf_{x} \quad f(x). \tag{5.1.1}$$

This problem can be expressed as a convex programming problem in $R^n \times R^m$:

$$\inf_{(x,z)} \quad \frac{1}{2}\langle z, z\rangle \tag{5.1.2}$$

$$\text{subject to:} \quad F(x) - z \leq 0.$$

The infimum of this problem (5.1.2) exists, since the constraints are feasible and the objective function is bounded below. If a solution (x, z) exists, then z is unique since the constraints are convex and the objective function is strictly convex . Let $\mathcal{J}(x)$ be the Jacobian of $F(x)$. Then the KKT conditions for any solution (x, z) of (5.1.2) are:

$$\begin{aligned} \mathcal{J}^T(x)z &= 0 \in R^n, \\ z &\geq 0 \in R^m, \\ F(x) - z &\leq 0 \in R^m, \\ \langle z, F(x) - z \rangle &= 0. \end{aligned} \tag{5.1.3}$$

In section 3.2, numerical stability (based on the definition of Robinson[1975]) of the solution set of the least square formulation for solving systems of linear inequalities was established. Robinson[1976] treats the case of systems of inequalties of nonlinear functions and shows an equivalence between regular systems and stable solution sets. For the case of a system of inequalities of convex C^2 functions defined in (5.1.3), the formulation as a least square problem also yields a system of inequalities which are regular and hence the solution set is stable.

5.2 Approximating Properties of Taylor Series Using Generalized Differential Constructs

Since we assume F is convex and C^2 and hence finite, then $F_+(x)$ is Lipschitz continuous of some rank γ in a neighborhood of x, and then there is a well-defined generalized Jacobian set of $F_+(x)$. Since the gradient of $f(x)$ exists and is Lipschitz continuous on lower level sets, then there is a well-defined generalized Hessian set of $f(x)$, comparable to the linear case in Section 3. We use the same notations for these constructs in this section, as in Section 3. Denote the Jacobian of F at x as $\mathcal{J}(x)$. Using the definition of the generalized Jacobian, similar as in Section 3, the possible choices of generalized Jacobian $J(x)$ of $F_+(x)$ are:

$$J(i,j) = \begin{cases} 0 & \text{if } F_i(x) < 0, \\ \mathcal{J}(i,j) & \text{if } F_i(x) > 0, \\ \lambda_i \mathcal{J}(i,j) & \text{if } F_i(x) = 0, \text{ where } \lambda_i \in [0,1]. \end{cases}$$

Proposition 5.2.1. *Existence of Indentification Neighborhood*

Let x_* be given and let

$$\begin{aligned} I^0(x_*) &:= \{i | F_i(x_*) = 0\}, \\ I^+(x_*) &:= \{i | F_i(x_*) > 0\}, \\ I^-(x_*) &:= \{i | F_i(x_*) < 0\}. \end{aligned}$$

Then there exists a neighborhood $N(x_*, r)$ of x_*, such that for $x \in N(x_*, r)$,

$$\begin{aligned} I^0(x) &\subset I^0(x_*), \\ I^+(x_*) &= I^+(x) \backslash I^0(x_*), \\ I^-(x_*) &= I^-(x) \backslash I^0(x_*). \end{aligned} \tag{5.2.1}$$

Proof. Define $C_i := \{x | F_i(x) \leq 0\}$, and note that since each F_i is convex and C^2 and hence finite, then each C_i is closed and convex. If $i \in I^-(x_*)$, then $F_i(x_*) < 0$, and since each F_i is continuous there exists a neighborhood $N(x_*, r_i)$ about x_*, such that for $x \in N(x_*, r_i)$ implies $F_i(x) < 0$. If $i \in I^+(x_*)$, then

$F_i(x_*) > 0$ and $x_* \notin C_i$ which is a closed convex set. Thus, there exists a neighborhood $N(x_*, r_i)$ about x_*, such that for $x \in N(x_*, r_i)$ implies $F_i(x) > 0$.

Choose $r > 0$ so that,

$$r < \min\{r_i\}.$$

If $x \in N(x_*, r)$, then

$$I^0(x) \subset I^0(x_*),$$

$$I^+(x_*) = I^+(x) \backslash I^0(x_*),$$

$$I^-(x_*) = I^-(x) \backslash I^0(x_*).$$

||

Proposition 5.2.1 shows the existence of an identification neighborhood holds for the C^2 convex functions. In the linear case, one can perfectly model $(Ax_* - b)_+$ from points in an identification neighborhood using the first order Taylor series based on any choice of generalized Jacobian. For the case of systems of inequalities of convex C^2 functions, the proposition below shows that the generalized Jacobian model attains first order approximation of $F_+(x_*)$ from any point x within this Proposition 5.2.1 neighborhood of a point x_*

Proposition 5.2.2. *First Order Approximation Using a Generalized Jacobian Model*

Let x_* be given and let $N(x_*, r)$ be an identification neighborhood which satisfies (5.2.1) in Proposition 5.2.1. Let $x \in N(x_*, r)$, and let $p := x_* - x$. Since each F_i is C^2 convex, denote γ to be the Lipschitz constant of $\mathcal{J}F$ on $N(x_*, r)$. Then the following holds:

1. For any $J(x) \in \partial F_+(x)$,

$$F_+(x + p) - F_+(x) - J(x)p = o(\|p\|),$$

and specifically,

$$\|F_+(x + p) - F_+(x) - J(x)p\| \leq \frac{\gamma}{2}\|p\|^2.$$

2. For $x \in N(x_*, r)$, any generalized Jacobian $J(x_*) \in \partial F_+(x_*)$ satisfies:

$$\|F_+(x) - F_+(x_*) - J(x_*)(-p)\| \leq \|\mathcal{J}_{I^0(x_*)}\| \, \|p\| + \frac{\gamma}{2}\|p\|^2.$$

3. For a particular $x \in N(x_*, r)$, there is at least one generalized Jacobian $J(x_*) \in \partial F_+(x_*)$, which satisfies for this particular x:

$$\|F_+(x) - F_+(x_*) - J(x_*)(-p)\| \leq \frac{\gamma}{2}\|p\|^2.$$

Proof. For 1, since $N(x_*, r)$ is bounded, and F is convex and C^2, and γ is the given Lipschitz constant of $\mathcal{J}(x)$, then the Jacobian of F on $N(x_*, r)$ satisfies, by Proposition 3.3.3,

$$\|F(x + p) - F(x) - \mathcal{J}(x)p\| \leq \frac{\gamma}{2}\|p\|^2.$$

Note that $x + p = x_*$. Consider any i:

1.1 If $i \notin I^0(x_*)$, so $i \notin I^0(x)$ and thus, $i \in I^-(x) \cup I^+(x)$. In these cases, $(J(x))_i$ is unique.

1.2 If $i \in I^0(x_*)$ and $i \in I^-(x) \cup I^+(x)$, then, as above, $(J(x))_i$ is unique. If $i \in I^0(x)$, then $F_i(x) = 0$ and since F_i is convex, then $F_i(x+\alpha p) \leq 0$, for any $\alpha \in [0,1]$. Since $F_i(x_*) = 0 = F_i(x)$ can be approximated first order about x, by using $(\mathcal{J}(x))_i$, then choosing $(J(x))_i = \lambda(\mathcal{J}(x))_i$, for any $\lambda \in [0,1]$, gives just as good or an even better approximation. Specifically, from the Lipschitzian property of $\mathcal{J}(x)$, for some γ_i:

$$\|F_i(x_*) - F_i(x) - (\mathcal{J}(x))_i p\| \leq \frac{\gamma_i}{2}\|p\|^2.$$

This means that

$$\begin{aligned}
&\|(F_+(x_*))_i - (F_+(x))_i - \lambda(\mathcal{J}(x))_i p\| \\
&= \| - \lambda(\mathcal{J}(x))_i p\| \\
&\leq \lambda \frac{\gamma_i}{2}\|p\|^2 \\
&\leq \frac{\gamma_i}{2}\|p\|^2.
\end{aligned}$$

In either case 1.1 or 1.2, then

$$\|F_+(x+p) - F_+(x) - J(x)p\| \leq \frac{\gamma}{2}\|p\|^2.$$

For 2, since $\mathcal{J}F$ is Lipschitz continuous of rank γ, and since all the error in approximating $F_+(x)$ using a first order approximation at x_* is contained in the components with indices $i \in I^0(x_*)$, then

$$\|F_+(x) - F_+(x_*) - J(x_*)(-p)\| \leq \|\mathcal{J}_{I^0(x_*)}\| \, \|p\| + \frac{\gamma}{2}\|p\|^2.$$

For 3, let a particular $x \in N(x_*, r)$ be given and let

$$\begin{aligned}
I^0 &:= I^0(x_*) \backslash I^0(x), \\
I^+ &:= I^+(x) \cap I^0(x_*), \\
I^- &:= I^-(x) \cap I^0(x_*).
\end{aligned}$$

For $i \in I^+$, choose $(J(x_*))_i := (\mathcal{J}(x_*))_i$; for $i \in I^-$, choose $(J(x_*))_i := 0$; and for $i \in I^0$, choose $(J(x_*))_i := \lambda(\mathcal{J}(x_*))_i$, where $\lambda \in [0,1]$. Then this particular choice of $J(x_*) \in \partial F_+(x_*)$ satisfies for this particular x, by 1 and 2 above, and direct calculation,

$$\|F_+(x) - F_+(x_*) - J(x_*)(-p)\| \leq \frac{\gamma}{2}\|p\|^2.$$

||

The gradient of $f(x)$ is well-defined on R^n. $\nabla f(x) = J^T(x)F_+(x)$, for any choice of $J(x) \in \partial F_+(x)$. ∇f is continuous, being the gradient of a differentiable convex function. Since f is convex, bounded below and continuously differentiable, then the gradient of f is Lipschitz continuous (order 1) of some rank γ on lower level sets. If $M > 0$ and $S := \{x | f(x) \leq M\}$, then there exists $\gamma > 0$, such that for $u, v \in S$,

$$\|\nabla f(u) - \nabla f(v)\| \leq \gamma\|u - v\|. \tag{5.2.2}$$

Therefore, the generalized Hessian set of $\nabla f(x)$, denoted $\partial\nabla f(x)$, is well-defined on S. Selecting a generalized Hessian for f is done similarly as in section 3. Let Γ_i denote the Hessian of F_i. One can calculate the generalized Jacobian of $\nabla f(x) = J^T(x)F_+(x)$, using the definition. Let Ω be the set of points where ∇f is not differentiable. Calculating any limit H_k as $y_l \to x$:

$$\begin{aligned} H_k &= \lim_{y_l \to x} \nabla^2 f(y_l) \text{ where } y_l \notin \Omega \\ &= \lim_{y_l \to x} \mathcal{J}(\mathcal{J}(y_l)^T)F_+(y_l) + \mathcal{J}(y_l)^T\mathcal{J}(y_l) \text{ where } y_l \notin \Omega \\ &= \lim_{y_l \to x} \sum_{i \in I^+(y_l)} \Gamma_i(y_l)F_i(y_l) + \mathcal{J}(y_l)^T\mathcal{J}(y_l) \text{ where } y_l \notin \Omega, \end{aligned}$$

where $I^+(y_l) = \{i | F_i(y_l) > 0\}$. Using the same type of argument in Section 3, one can show the choices of generalized Hessians depend only on the choice of generalized Jacobian, $J(x) \in \partial F_+(x)$,

$$H(x) \in \partial(J^T(x)F_+(x)) \iff H(x) = J^T(x)J(x) + \sum_{i \in I^+(x)} \Gamma_i(x)F_i(x) \text{ for some } J(x) \in \partial F_+(x). \quad (5.2.3)$$

Proposition 3.3.8 assures that the generalized Hessian model (based on any generalized Hessian at x) attains perfect approximation of the linear inequalities least square function at x_* from any point x in an identification neighborhood of x_*. Proposition 5.2.3, below, shows that the generalized Hessian model (based on any generalized Hessian at x) attains second order approximation of $f(x_*)$, the least square function of inequalities of convex C^2 functions, from any point x in an identification neighborhood of x_*.

Proposition 5.2.3. *Second Order Approximation Using a Generalized Hessian Model*

Let x_* be given and let $N(x_*, r)$ be an identification neighborhood which satisfies (5.2.1) in Proposition 5.2.1. Since F is C^2 convex on $N(x_*, r)$, let γ be the order 1 Lipschitz constant for $\nabla^2(\frac{1}{2}\langle F(x), F(x)\rangle)$. Let $x \in N(x_*, r)$ and let $p := x_* - x$. Then for any $J(x) \in \partial F_+(x)$ which determines some $H(x) \in \partial\nabla f(x)$,

$$f(x+p) - f(x) - \langle p, \nabla f(x)\rangle + \frac{1}{2}\langle p, H(x)p\rangle = o\|p\|^2,$$

or specifically, for the Lipschitz constant, γ, of $\nabla^2(\frac{1}{2}\langle F(x), F(x)\rangle)$:

$$|f(x+p) - f(x) - \langle p, \nabla f(x)\rangle - \frac{1}{2}\langle p, H(x)p\rangle| \leq \frac{\gamma}{6}\|p\|^3.$$

Proof. The proof is similar to the proof of Proposition 5.2.2. Note that $x + p = x_*$. Let γ be an order 1 Lipschitz constant of $\nabla^2(\frac{1}{2}\langle F(x), F(x)\rangle)$ on $N(x_*, r)$, which exists xince $N(x_*, r)$ is bounded and F is convex C^2. Consider any i:

1. If $i \notin I^0(x_*)$, then $i \notin I^0(x)$. If $i \in I^-(x) \cup I^+(x)$, then $(J(x))_i$ is unique.
2. If $i \in I^0(x_*)$. If $i \in I^-(x) \cup I^+(x)$, then $(J_2(x))_i$ is unique. If $i \in I^0(x)$, then there are other possible generalized Jacobians, $J(x) \in \partial F_+(x)$. Notice that the Hessian term cancels out, since $F_i(x) = 0$. Since $\frac{1}{2}\langle F(x+p), F(x+p)\rangle$ can be approximated second order by the second order Taylor series about x using

the Hessian model with $\nabla^2 \frac{1}{2}\langle F(x), F(x)\rangle$, then choosing $(J(x))_i = \lambda(\mathcal{J}(x))_i$, for any $\lambda \in [0,1]$, gives just as good or an even better approximation, as in the proof of Proposition 5.2.2.

In either case, then

$$|f(x+p) - f(x) - \langle p, \nabla f(x)\rangle - \frac{1}{2}\langle p, H(x)p\rangle| \leq \frac{\gamma}{6}\|p\|^3.$$

||

From (5.2.3), any generalized Hessian $H(x)$ of f at x can be written as

$$\begin{aligned} H(x) &:= J^T(x)J(x) + Q(x) \\ \text{where } Q(x) &:= \sum_{i\in I^+(x)} \Gamma_i(x)F_i(x). \end{aligned} \tag{5.2.4}$$

Lemma 2.

Let x_* be given. Then there is some $\sigma > 0$ and there is a neighborhood $N(x_*, r)$ of x_*, such that for $x \in N(x_*, r)$, given any $J(x) \in \partial F_+(x)$ and any $J(x_*) \in \partial F_+(x_*)$,

$$\|(J(x) - J(x_*))^T F_+(x_*)\| \leq \sigma\|x - x_*\|.$$

Further, there is some $\beta > 0$ such that for any $x \in N(x_*, r)$, any $J(x) \in \partial F_+(x)$, and any $J(x_*) \in \partial F_+(x_*)$,

$$\|(J(x) - J(x_*))^T F_+(x_*) - Q(x_*)(x - x_*)\| \leq \beta\|x - x_*\|^2.$$

Proof. For the first part, let $I^+(x_*) := \{i|\,(F(x_*))_i > 0\}$. If $F(x_*) \leq 0$, then the lemma is satisfied. so consider the case where $I^+(x_*) \neq \emptyset$. Form a neighborhood, where $r > 0$ is chosen so that if $x \in N(x_*, r)$, then

$$I^+(x_*) \subset I^+(x) := \{i|\,(F(x))_i > 0\}.$$

For indices $i \in I^+(x_*)$, then $(J(x))_i$ is unique and $(J(x_*))_i$ is unique. Since F is convex C^2, then $\mathcal{J}F(x)$ is Lipschitz continuous, of some rank γ of order 1 on $N(x_*, r)$. Then one can approximate:

$$\|(J(x) - J(x_*))^T F_+(x_*)\| = \|\sum_{i\in I^+(x_*)} (J_i(x) - J_i(x_*))F_i(x_*)\| \leq \gamma\|F_+(x_*)\|\,\|x - x_*\| \leq \sigma\|x - x_*\|,$$

for any $\sigma \geq \gamma\|F_+(x_*)\|$.

For the second part, since F is C^2 convex, let γ_Q be the order 1 Lipschitz constant of $Q(x)$ on $N(x_*, r)$. This implies there exist γ_i for $i \in I^+(x_*)$ and $x \in N(x_*, r)$, such that

$$\|(J(x) - J(x_*))_i - \Gamma_i(x_*)(x - x_*)\| \leq \gamma_i\|x - x_*\|^2.$$

It follows for any $x \in N(x_*, r)$, any $J(x) \in \partial F_+(x)$, $J(x_*) \in \partial F_+(x_*)$, then

$$\begin{aligned} \|(J(x) - J(x_*)^T F_+(x_*) - Q(x_*)(x - x_*)\| &= \|\sum_{i\in I^+(x_*)} F_i(x_*)((J(x) - J(x_*))_i - \Gamma_i(x_*)(x - x_*))\| \\ &\leq \|F_+(x_*)\|\gamma_Q\|x - x_*\|^2 \\ &\leq \beta\|x - x_*\|^2, \end{aligned}$$

since for $i \in I^+(x_*)$, $(J(x))_i$ is unique and $(J(x_*))_i$ is unique. ||

Lemma 3, below, shows that the first order Taylor series of ∇f (about a point x in an identification neighborhood of a point x_*) based on a generalized Hessian of $\partial\nabla f(x)$ attains first order approximation. The lemma also shows for each particular point x, the existence of at least one generalized Hessian $H_x^* \in \partial\nabla f(x_*)$ for which the Taylor series about x_* models $\nabla f(x)$ first order. The notation $H_x^* \in \partial\nabla f(x_*)$ is used since the Taylor series about x_* using this particular generalized Hessian H_x^* models $\nabla f(x)$.

Lemma 3. *First Order Approximation of* ∇f.

Let x_* be given. Let $N(x_*, r)$ be a Proposition 5.2.1 identification neighborhood of x_*. Since F is C^2 convex, let γ be an order 1 Lipschitz constant of $\nabla^2(\frac{1}{2}\langle F(x), F(x)\rangle)$ on $N(x_*, r)$. Then for $x \in N(x_*, r)$ and for any $J(x) \in \partial F_+(x)$ which determines some $H(x) \in \partial\nabla f(x)$,

$$\|\nabla f(x_*) - \nabla f(x) - H(x)(x_* - x)\| = o\|x_* - x\|,$$

or specifically, for the Lipschitz constant γ:

$$\|\nabla f(x_*) - \nabla f(x) - H(x)(x_* - x)\| \leq \frac{\gamma}{2}\|x_* - x\|^2.$$

Additionally, for each particular $x \in N(x_*, r)$, there exists a generalized Hessian, $H_x^* \in \partial\nabla f(x_*)$, such that

$$\|\nabla f(x) - \nabla f(x_*) - H_x^*(x - x_*)\| \leq \frac{\gamma}{2}\|x - x_*\|^2.$$

Proof. Let $p = x_* - x$. For the first conclusion, use the same argument as for Proposition 5.2.2. Since F is convex C^2 on $N(x_*, r)$, then $\nabla^2\frac{1}{2}\langle F(x), F(x)\rangle$ is Lipschitz continuous (order 1) of some given rank γ. Then

$$\|\mathcal{J}(x+p)^T F(x+p) - \mathcal{J}(x)^T F(x) - \mathcal{J}^T(x)\mathcal{J}(x)p - \sum_i F_i(x)\Gamma_i(x)p\| \leq \frac{\gamma}{2}\|p\|^2.$$

Now, for each i consider the same cases as in Proposition 5.2.2. The only case where there is not a unique choice of generalized Jacobian, $J(x) \in \partial F_+(x)$, is for $i \in I^0(x_*) \cap I^0(x)$, where again, any choice of generalized Jacobian gives a even better or just as good an approximation, as in Proposition 5.2.2. Therefore, the first conclusion follows.

The second result uses the same argument as in Proposition 5.2.2. ||

An approximating property of Lipschitz continuous Jacobians follows:

Proposition 5.2.4.

Let $G : \Re^n \mapsto \Re^m$ be C^1. Let x_* be given, and let the Jacobian of G, denoted $\mathcal{J}G$, be Lipschitz continuous of order γ on some neighborhood $N(x_*, r)$. Let $u, v \in N(x_*, r)$. Then,

$$\|G(v) - G(u) - \mathcal{J}(x_*)(v - u)\| \leq \frac{\gamma}{2}(\|v - x_*\| + \|u - x_*\|)\|v - u\|.$$

Proof.

$$\begin{aligned}
\|G(v) - G(u) - \mathcal{J}(x_*)(v-u)\| &= \| \left(\int_0^1 \mathcal{J}(u + t(v-u))\, dt \right)(v-u) - \mathcal{J}(x_*)(v-u)\| \\
&\le \int_0^1 \|\mathcal{J}(u + t(v-u)) - \mathcal{J}(x_*)\|\, dt\, \|(v-u)\| \\
&\le \int_0^1 \gamma \|(1-t)u + tv - x_*\|\, dt\, \|(v-u)\| \\
&= \int_0^1 \gamma \|(1-t)(u - x_*) + t(v - x_*)\|\, dt\, \|(v-u)\| \\
&\le \gamma \left(\int_0^1 (1-t)\, dt\, \|(u-x_*)\| + \int_0^1 t\, dt\, \|(v-x_*)\| \right) \|(v-u)\| \\
&= \frac{\gamma}{2}(\|v - x_*\| + \|u - x_*\|)\, \|v-u\|.
\end{aligned}$$

||

The result below shows one more approximating property, relating other points in the identification neighborhood of a point.

Lemma 4.

Let x_* be given. Let $N(x_*, r)$ be a Proposition 5.2.1 identification neighborhood of x_*, and let $u, v \in N(x_*, r)$. Then there is some $\sigma \ge 0$, such that any $J(v) \in \partial F(v)$, any $J(u) \in \partial F(u)$ and $Q(x_*) = \sum_{I^+(x_*)} \Gamma_i(x_*) F_i(x_*)$ satisfy:

$$\|(J(v) - J(u))^T F_+(x_*) - Q(x_*)(v-u)\| \le \sigma(\|v - x_*\| + \|u - x_*\|)\|v-u\|$$

Proof. Let $N(x_*, r)$ be an identification neighborhood of x_*. Let γ_i be the order 1 Lipschitz constant of Γ_i on $N(x_*, r)$, and let $u, v \in N(x_*, r)$ and $J(v) \in \partial F(v)$, $J(u) \in \partial F(u)$. Then

$$\|(J(v) - J(u))^T F_+(x*) - Q(x_*)(v-u)\| = \| \sum_{i \in I^+(x_*)} F_i^+(x_*)((J(v) - J(u))_i - \Gamma_i(x_*)(v-u))\|.$$

Since $I^+(x_*) \subset I^+(u)$ and $I^+(x_*) \subset I^+(u)$, then if $i \in I^+(x_*)$, then $i \in I^+(v)$ and $i \in I^+(u)$, and $(J(v))_i = (\mathcal{J}(v))_i$, $(J(u))_i = (\mathcal{J}(u))_i$. By Proposition 5.2.4 above, for each $i \in I^+(x_*)$, then

$$\|(J(v) - J(u))_i - \Gamma_i(x_*)(v-u)\| \le \frac{\gamma_i}{2}(\|v - x_*\| + \|u - x_*\|)\|v-u\|,$$

and then

$$\begin{aligned}
&\| \sum_{i \in I^+(x_*)} F_i^+(x_*)((J(v) - J(u))_i - \Gamma_i(x_*)(v-u))\| \le \\
&\sum_{i \in I^+(x_*)} F_i^+(x_*) \frac{\gamma_i}{2}(\|v - x_*\| + \|u - x_*\|)\|v-u\| \le \\
&\sigma(\|v - x_*\| + \|u - x_*\|)\|v-u\|,
\end{aligned}$$

where $\sigma = \| \sum_{i \in I^+(x_*)} \frac{\gamma_i}{2} F_i^+(x_*)\|$.

||

6 Generalization of Algorithms for Solving a System of Inequalities of Convex C^2 Functions

In this section, we develop algorithms for solving the system of inequalities of C^2 convex functions in a least square sense and analyze convergence properties of these algorithms. Refer to the algorithms described in section 1.5 for the solution of least square problem of systems of equalities. We extend these algorithms for the case of inequalities of convex C^2 functions using generalized differential constructs. In section 6.1, local convergence results are addressed. We extend Newton's method and show q-quadratic local convergence. We compare this method with the generalized Newton's algorithm for nonsmooth functions developed by Robinson[1990] which he bases on a point-based approximation. We develop and analyze a quasi-Newton symmetric secant update in this section and define a bounded deterioration property for secant updates using generalized differential constructs. By assuming this property holds for the secant updates, we show local q-linear convergence. The extensions of Gauss-Newton and Levenberg-Marquardt methods for the case of systems of convex C^2 inequalities are discussed, as well as the local convergence results. In related work, Burke[1983], and Burke and Han[1986] describe a method for solving this type of system of inequalities, using the distance function as a penalty type objective function and projections to derive search directions. They present global convergence results for their search direction using an Armijo line search method. Their local second order convergence result for the non-linear functions depends on the Mangasarian Fromowitz constraint qualification (MFCQ) condition, which is discussed with the results in this section.

In section 6.2, we discuss global convergence of the Armijo line search and trust region method applied to solving a system of inequalities of convex C^2 functions and equations of non-linear C^2 functions in least square sense.

Consider the problem, similar to (3.1.2), with the function F_2, now being convex in each dimension, and C^2. That is:

$$F_1 : \Re^n \mapsto \Re^{m1} \quad \text{where } F_1 \text{ is } C^2,$$

$$F_2 : \Re^n \mapsto \Re^{m2} \quad \text{where } F_2 \text{ is } C^2 \text{ and component-wise convex,}$$

and we wish to solve:

$$\inf f(x) \quad \text{where } f(x) := f_1(x) + f_2(x) := \frac{1}{2}\langle F_1(x), F_1(x)\rangle + \frac{1}{2}\langle F_2^+(x), F_2^+(x)\rangle. \tag{6.1}$$

Denote Γ_1^i to be the Hessian of $(F_1)_i$ and Γ_2^i to be the Hessian of $(F_2)_i$. Denote $J_1(x)$ to be the Jacobian of $F_1(x)$, and $J_2(x)$ to be a generalized Jacobian of F_2^+. Then a generalized Hessian of f is:

$$H(x) := H_1(x) + H_2(x) = J_1^T(x)J_1(x) + \sum_{i=1}^{m1} \Gamma_1^i(x)(F_1(x))_i + J_2^T(x)J_2(x) + \sum_{i\in I^+(x)} \Gamma_2^i(x)(F_2(x))_i.$$

6.1 Local Convergence Results

Local convergence results use the unscaled search direction, as in section 4.1.

6.1.1 Newton-like Method for Solving $\nabla f(x) = 0$: Define the generalized Hessian Newton Method for minimization, where H_k denotes a generalized Hessian at iterate x_k. Then d_k solves:

$$H_k d_k = -\nabla f(x_k),$$

where $x_{k+1} := x_k + d_k$. A solution d_k exists and is unique if H_k is invertible.

Theorem 5. *Newton-like Method for Solving Convex Inequalities and Non-linear Equations in a Least Square Sense.*

Let f be given as in (6.1), where F_1 is C^2 and F_2 is C^2 convex. Assume there exists x_*, such that $\nabla f(x_*) = 0$.

1. Let $r > 0$, and let $N(x_*, r)$ be a neighborhood on which $\nabla^2 f_1$ is Lipschitz continuous of order 1 for some rank. $\gamma_1 > 0$.
2. Let some $\beta > 0$ be such that any generalized Hessian $H(x_*) \in \partial\nabla f(x_*)$ is invertible and satisfies:

$$\|H^{-1}(x_*)\| \leq \frac{\beta}{2}.$$

Then there exists $\epsilon > 0$, such that if $x_0 \in N(x_*, \epsilon)$, then the sequence of iterates, $x_{k+1} := x_k - H^{-1}(x_k)\nabla f(x_k)$ are well defined and converge to x_* and satisfy for some $\gamma > 0$,

$$\|x_{k+1} - x_*\| \leq \beta\gamma\|x_k - x_*\|^2.$$

Proof. Since, by hypothesis, all the generalized Hessians $H(x_*) \in \partial\nabla f(x_*)$ are invertible, Proposition 1.3.1. assures a neighborhood of invertibility about x_*. By choosing ϵ sufficiently small, for any $x \in N(x_*, \epsilon)$, and any $H(x) \in \partial\nabla f(x)$, then $H^{-1}(x)$ exists and

$$\|H^{-1}(x)\| \leq \beta.$$

By Lemma 3, since F_2 is convex and C^2, and choosing ϵ sufficiently smaller if necessary so that $N(x_*, \epsilon)$ is an identification neighborhood of x_*, then for some γ_2, for $x \in N(x_*, \epsilon)$ and any generalized Hessian of $f_2(x)$, $H_2(x)$, it follows that

$$\|\nabla f_2(x_*) - \nabla f_2(x) - H_2(x)(x_* - x)\| \leq \frac{\gamma_2}{2}\|x_* - x\|^2.$$

Given 1, that $\nabla^2 f_1$ is Lipschitz continuous of some rank γ_1 on $N(x_*, r)$, then for $\gamma = \gamma_1 + \gamma_2 > 0$, choosing $\epsilon \leq r$, then for all $x \in N(x_*, \epsilon)$:

$$\|\nabla f(x_*) - \nabla f(x) - H(x)(x_* - x)\| \leq \gamma\|x_* - x\|^2.$$

Choose ϵ smaller, if necessary, to satisfy:

$$\epsilon < \min\left(1, \frac{1}{\beta\gamma}\right).$$

We show the case for $k = 0$. Since $H^{-1}(x)$ exists on $N(x_*, \epsilon)$, then x_1 is well-defined and satisfies, for any generalized Hessian of $H(x_0)$:

$$\begin{aligned} x_1 - x_* &= x_0 - x_* - H^{-1}(x_0)\nabla f(x_0) \\ &= x_0 - x_* - H^{-1}(x_0)[\nabla f(x_0) - \nabla f(x_*)] \\ &= H^{-1}(x_0)[\nabla f(x_*) - \nabla f(x_0) - H(x_0)(x_* - x_0)]. \end{aligned}$$

Using the result above, then

$$\begin{aligned} \|x_1 - x_0\| &\leq \|H^{-1}(x_0)\| \, \|[\nabla f(x_*) - \nabla f(x_0) - H(x_0)(x_* - x_0)]\| \\ &\leq \beta\gamma\|x_0 - x_*\|^2. \end{aligned}$$

Using induction for the general k and $k+1$ is accomplished the same way. ||

This result establishes q-quadratic local convergence of this Newton-like method based on generalized Hessians. Robinson[1990] extended Newton's method for nonsmooth functions having a point based approximation satisfying second order approximation. For our special least square function, using a linearization about x_k based on a generalized Hessian at x_k to model $\nabla f(x_*)$ assures a second order point based second order approximation. Robinson's point based approximation is not required to be linear and thus, rather than using the Banach lemma for invertibility, he develops an implicit function property of invertibility which handles the non-linear point based approximation. Our result is specialized in that we define a specific linear point based approximation. The result in Theorem 5 closely parallels the traditional proof for the smooth case. Robinson showed how to use a point based approximation to get a convergent Newton's method, while this work shows that a system of linear equalities naturally gives rise to a point based approximation.

6.1.2 Adaptation of Symmetric Secant Update Quasi-Newton Direction: A problem with using Newton's method for finding search directions is that the Hessian or generalized Hessian may not be positive definite, and possibly not invertible, so d may not be a descent direction. However, even if the generalized Hessians are all positive definite in a neighborhood of a stationary point, the difficulty in computing the Hessian at each iterate is another reason that practical algorithms use a quasi-Newton method. This approach uses Hessian approximations, avoiding computational steps to calculate the full Hessian. We extend the special symmetric secant update method for least square nonlinear equations to solve for a zero of $\nabla f(x)$. The method uses the available generalized Jacobians, $J(x)$, and uses a secant update, B_k, to approximate the second order term:

$$B_k \approx Q(x_k) = \sum_{i}^{m1}(F_1(x_k))_i\Gamma_1^i(x_k) + \sum_{i\in I^+(x_k)}(F_2(x_k))_i\Gamma_2^i(x_k).$$

Denote

$$F(x) = \begin{pmatrix} F_1(x) \\ F_2^+(x) \end{pmatrix}.$$

An algorithm using the symmetric secant update for generalized Jacobians is as follows: Solve for d_k

$(J^T(x_k)J(x_k) + B_k)d_k = -\nabla f(x_k)$, and set:

$$\begin{aligned} x_{k+1} &:= x_k + d_k, \\ y_k &:= \nabla f(x_{k+1}) - \nabla f(x_k), \\ y_k^{\#} &:= J^T(x_{k+1})F(x_{k+1}) - J^T(x_k)F(x_{k+1}), \\ B_{k+1} &:= B_k + \frac{(y_k^{\#} - B_k d_k)d_k^T + d_k(y_k^{\#} - B_k d_k)^T}{\langle y_k, d_k \rangle} - \frac{\langle y_k^{\#} - B_k d_k, d_k \rangle y_k d_k^T}{\langle y_k, d_k \rangle^2}. \end{aligned} \tag{6.1.2.1}$$

The updated B_{k+1} satisfies the quasi-Newton condition:

$$B_{k+1} d_k = y_k^{\#}.$$

Dennis, Gay, and Welsch[1981] developed this special symmetric secant update for a least square solution of non-linear equations. For the smooth case, they show this update solves the following least change problem:

$$\min_{B_{k+1} \in \Re^{n \times n}} \quad \|(T^T)^{-1}(B_{k+1} - B_k)T^{-1}\|_F$$

$$\text{subject to} \quad B_{k+1} - B_k \text{ symmetric}, \quad B_{k+1}d_k = y_k^{\#},$$

where $T \in \Re^{n \times n}$, a weighting matrix, is non-singular and satisfies $TT^T d_k = y_k$.

Since this method is specially tailored for the least square solution of non-linear equations, it is an appropriate quasi-Newton method to extend to the least square solution of inequalities of non-linear convex C^2 functions. This update is called a rescaled symmetric secant update, because of the weighting matrix T. This update and the unscaled update are analyzed in Dennis and Walker[1981] for the smooth case. Local convergence analysis of quasi-Newton methods for generalized differential constructs requires that we define a bounded deterioration property for the updates, B_k.

Definition. *Bounded Deterioration Property for Generalized Hessians.*

Let $N(x_*, r)$ be an identification neighborhood of a stationary point x_*, with iterates $x_k \in N(x_*, r)$. Define $\bar{J}(x_k) \in \partial F_+(x_k)$, by $(\bar{J}_2(x_k))_i = 0$, for $i \in I^0(x_k)$. (Note that for indices $i \in I^+(x_k) \cup I^-(x_k)$, that $J_2(x_k))_i$ is unique.) Define $J_k^* \in \partial F_+(x_*)$ to be the generalized Jacobian which models $x_k - x_*$, as defined in Proposition 5.2.2, and choose $(J_2^*)_k^i = 0$ for indices $i \in I^0(x_*)$. (Again, note that for indices $i \in I^+(x_*) \cup I^-(x_*)$, that $(J_2^*)_k^i$ is unique.) Say the updates B_k, such as in (6.1.2.1), satisfy the *bounded deterioration property* if there exists an α, such that for all k,

$$\|\bar{J}_{k+1}^T \bar{J}_{k+1} - (J_{k+1}^*)^T J_{k+1}^* + B_{k+1} - Q_*\| \le \|\bar{J}_k^T \bar{J}_k - (J_k^*)^T J_k^* + B_k - Q_*\| + \alpha \max(\|x_{k+1} - x_*\|, \|x_k - x_*\|). \tag{6.1.2.2}$$

The particular $\bar{J}_k \in \partial\nabla f(x_k)$ and $J_k^* \in \partial\nabla f(x_*)$ that are chosen allow consistent comparison of (6.1.2.2) at each iteration k. Since in an identification neighborhood, $I^0(x_k) \subset I^0(x_*)$, and since J_k^* models $x_k - x_*$, the distance between these two generalized Jacobians is the distance between the compact convex sets, $\partial F(x_k)$ and $\partial F(x_*)$. One could choose other generalized Jacobians in the definition of the bounded deterioration property if this selection of generalized Jacobian, $J(x_k) \in \partial F(x)$ is compared properly with the closest element $J(x_*) \in \partial F(x_*)$. By choosing a particular generalized Jacobian at each iteration, $(\bar{J}(x_k))_i = 0$, for $i \in I^0(x_k)$, and if the updates satisfy the bounded deterioration property, then q-linear convergence is assured in some identification neighborhood, which is also a neighborhood of invertibility.

Theorem 6. *Local Convergence of Symmetric Secant Method for Inequalities.*

Let f be given as in (6.1). Let F_1 be C^2 and F_2 be C^2 convex. Let x_* be given and assume $\nabla f(x_*) = 0$. Let $N(x_*, r)$ be an identification neighborhood of x_*. Assume $\nabla^2 f_1$ is Lipschitz continuous of some rank γ_1 and let γ_2 be the order 1 Lipshitz constant of $\nabla^2 \frac{1}{2}\langle F_2(x), F_2(x)\rangle$ on $N(x_*, r)$, where $\gamma := \gamma_1 + \gamma_2$. Assume that all generalized Hessians, $H(x_*) \in \partial\nabla f(x_*)$ are invertible and satisfy on $N(x_*, r)$, $\|H^{-1}(x_*)\| \leq \beta$, for some β. Suppose the sequence of iterates and updates (6.1.2.1), where the generalized Jacobians at x_k are chosen so that $(\bar{J}(x_k))_i = 0$, for $i \in I^0(x_k)$, satisfy the bounded deterioration property on $N(x_*, r)$ for some α. Then there exists positive constants ϵ and δ, such that if $\|x_0 - x_*\| \leq \epsilon$ and $\|\bar{J}_0^T \bar{J}_0 - (J_0^*)^T J_0^* + B_0 - Q_*\| \leq \delta$, then the sequence of symmetric secant update iterates generated by (6.1.2.1) are well-defined and converge q-superlinearly to x_*.

Proof. Choose ϵ small enough to insure a neighborhood of invertibility of x_*, choose δ to satisfy:

$$6\beta\delta \leq 1, \tag{6.1.2.3}$$

and, if necessary, choose ϵ sufficiently smaller to also satisfy:

$$3\alpha\epsilon \leq \delta, \text{ and } 3\gamma\epsilon \leq 2\delta. \tag{6.1.2.4}$$

We prove the two statements (6.1.2.5) and (6.1.2.6) below, by induction on k:

$$\|\bar{J}_k^T \bar{J}_k - (J_k^*)^T J_k^* + B_k - Q_*\| \leq (2 - 2^{-k})\delta, \tag{6.1.2.5}$$

where $J_k^* \in \partial F_+(x_*)$ is the generalized Jacobian of Proposition 5.2.2 which models $x_k - x_*$, and $\bar{J}_k \in \partial F_+(x_k)$ is the generalized Jacobian chosen so that $(\bar{J}_k)_i = 0$, for the indices $i \in I^0(x_k)$ as in the definition of the bounded deterioration property for generalized Hessians.

$$\|x_{k+1} - x_*\| \leq \frac{1}{2}\|x_k - x_*\|. \tag{6.1.2.6}$$

For $k = 0$, (6.1.2.5) is satisfied by assumption in the hypothesis. We show (6.1.2.6) in the induction step, which is comparable to the case for $k = 0$. Assume (6.1.2.5) and (6.1.2.6) hold for $k = 1, 2, \ldots, l - 1$.

For $k = l$, by the bounded deterioration property assumption, and the two induction hypothesis, then the selected generalized Jacobians, J_l^*, which models $x_l - x_*$, and $\bar{J}_l$ satisfy:

$$\|\bar{J}_l^T \bar{J}_l - (J_l^*)^T J_l^* + B_l - Q_*\| \leq (2 - 2^{-(l-1)})\delta + \alpha\|x_{l-1} - x_*\|. \qquad (6.1.2.7)$$

From (6.1.2.6) and given that $\|x_0 - x_*\| \leq \epsilon$, then

$$\|x_{l-1} - x_*\| \leq 2^{-(l-1)}\|x_0 - x_*\| \leq 2^{-(l-1)}\epsilon.$$

Substituting this into (6.1.2.7) and using (6.1.2.4) yields

$$\begin{aligned}\|\bar{J}_l^T \bar{J}_l - (J_l^*)^T J_l^* + B_l - Q_*\| &\leq (2 - 2^{-(l-1)})\delta + \alpha 2^{-(l-1)}\epsilon \\ &\leq (2 - 2^{-(l-1)} + 2^{-l})\delta = (2 - 2^{-l})\delta.\end{aligned}$$

which verifies (6.1.2.5). To verify (6.1.2.6), first show that $\bar{J}_l^T \bar{J}_l + B_l$ is invertible so that the iterates are well-defined. Given by hypothesis that any generalized Hessian in $\partial\nabla f(x_*)$ is invertible and satisfies $\|H^{-1}(x_*)\| \leq \beta$, from (6.1.2.5) and the induction hypothesis (6.1.2.7),

$$\|H^{-1}(x_*)(\bar{J}_l^T \bar{J}_l + B_l - H(x_*)\| \leq \|H^{-1}(x_*)\| \, \|(\bar{J}_l^T \bar{J}_l + B_l - H(x_*)\| \leq \beta(2 - 2^{-l})\delta < 2\beta\delta \leq \frac{1}{3}.$$

So, from the Banach lemma on invertibility of perturbed linear operators, $\bar{J}_l^T \bar{J}_l + B_l$ is invertible and satisfies:

$$\|(\bar{J}_l^T \bar{J}_l + B_l)^{-1}\| \leq \frac{\|H^{-1}(x_*)\|}{1 - \|H^{-1}(x_*)(\bar{J}_l^T \bar{J}_l + B_l - H(x_*))\|} \leq \frac{\beta}{1 - \frac{1}{3}} = \frac{3\beta}{2}. \qquad (6.1.2.8)$$

Thus, x_{l+1} is well defined and for the generalized Jacobian $J_l^* \in \partial F_+(x_*)$ which models $x_l - x_*$:

$$\|x_{l+1} - x_*\| \leq \|(\bar{J}_l^T \bar{J}_l + B_l)^{-1}\| \left[\| - \nabla f(x_*) + \nabla f(x_l) + H_l^*(x_* - x_l)\| + \|(\bar{J}_l^T \bar{J}_l + B_l - H_l^*\| \, \|(x_* - x_l)\|\right],$$

where $H_l^* := (J_l^*)^T J_l^* + Q_*$ is the generalized Hessian which models $x_l - x_*$. From Lemma 3,

$$\| - \nabla f(x_*) + \nabla f(x_l) + H_l^*(x_* - x_l)\| \leq \frac{\gamma\|x_l - x_*\|^2}{2}.$$

Substituting (6.1.2.5), (6.1.2.8), and from above:

$$\|x_{l+1} - x_*\| \leq \frac{3}{2}\beta\left[\frac{\gamma}{2}\|x_l - x_*\| + (2 - 2^{-l})\delta\right]\|x_l - x_*\|. \qquad (6.1.2.9)$$

From (6.1.2.6), the hypothesis that $\|x_0 - x_*\| \leq \epsilon$, and (6.1.2.4), then

$$\frac{\gamma\|x_l - x_*\|}{2} \leq 2^{-l+1}\gamma\epsilon \leq \frac{2^{-l}\delta}{3},$$

which substituted into (6.1.2.9), gives:

$$\begin{aligned}\|x_{l+1} - x_*\| &\leq \frac{3}{2}\beta\left[\frac{2^{-l}}{3} + 2 - 2^{-l}\right]\delta\|x_l - x_*\| \\ &\leq 3\beta\delta\|x_l - x_*\| \\ &\leq \frac{\|x_l - x_*\|}{2},\end{aligned}$$

with the final inequality coming from (6.1.2.4). This proves (6.1.2.6) for the induction step, l. Using these two induction results, it follows that the iterates converge q-linearly to x_*. ‖

Dennis and Walker[1981] show the conditions for the smooth case for the bounded deterioration property to be met. First of all, the set of matrices from which to choose the least change update, B_k, should include the matrix Q_*. Since Q_* is symmetric, positive semi-definite, the symmetric secant update (6.1.2.1) satisfies this condition. The other condition is whether there is some κ such that for all k, $y_k^{\#}$ satisfies:

$$\begin{aligned} \|y_k^{\#} - Q_*(x_{k+1} - x_k)\| &= \|(\bar{J}_{k+1} - \bar{J}_k)^T F_+(x_{k+1}) - Q_*(x_{k+1} - x_k)\| \\ &\leq \kappa \max\{\|x_k - x_*\|, \|x_{k+1} - x_*\|\}\|x_{k+1} - x_k\|. \end{aligned} \tag{6.1.2.10}$$

Notice that (6.1.2.10) is similar to the inequality in Lemma 4, which states the existence of $\sigma \geq 0$ such that for Q_*, J_{k+1}, J_k,

$$\|(J_{k+1} - J_k)^T F^+(x_*) - Q_*(x_{k+1} - x_k)\| \leq \sigma(\|x_k - x_*\| + \|x_{k+1} - x_*\|)\|x_{k+1} - x_k\|.$$

For the smooth case, where f is C^2, the authors show that the choice of

$$y_k^{\#} := \nabla f(x_{k+1}) - \nabla f(x_k) - \mathcal{J}^T(x_{k+1})\mathcal{J}(x_{k+1})(x_{k+1} - x_k)$$

always satisfies the bounded deterioration property. However, they cite that update (6.1.2.1) works just as well in test cases. Update (6.1.2.1) is used in NL2SOL algorithm, Dennis, Gay, and Walker[1981]. This algorithm is a hybrid algorithm which uses the model trust region method (Levenberg-Marquardt), with a zero matrix, B_0, unless the secant update gives a better match between predicted and actual reduction. The choice of comparison of generalized Jacobians and the global convergence of $x_k \to x_*$ may assure this condition (6.1.2.10) can be met in practice. For the smooth case, the authors show that NL2SOL performs better in test cases than the Levenberg-Marquardt algorithm, especially for the non-linear, large residual case.

Dennis and Walker[1981, Theorem 3.3] show for the smooth case that, if the iterates converge q-linearly, then the iterates converge q-superlinearly if and only if

$$\lim_{k \to \infty} \frac{\|(\bar{J}_k^T \bar{J}_k - (J_k^*)^T J_k^* + B_k - Q_*)(x_{k+1} - x_k)\|}{\|x_{k+1} - x_k\|} = 0 \tag{6.1.2.11}$$

6.1.3 Adaptation of Gauss-Newton for Convex Inequalities. If the residuals are small, it may be advantageous to use the Gauss-Newton method for the nonlinear inequality case, since the computation is somewhat simpler. We adapt the local convergence result for the case of inequalities of convex C^2 functions, combined with the non-linear equalities.

Theorem 7. *Local Convergence of Gauss-Newton for Solving Least Square Solution of System of Inequalities of Convex C^2 Functions and Non-linear Equalities.*

Let $F_1(x), F_2(x)$ be given as above and

$$f(x) := f_1(x) + f_2(x) := \frac{1}{2}\langle F_1(x), F_1(x)\rangle + \frac{1}{2}\langle F_2^+(x), F_2^+(x)\rangle,$$

and assume F_2 is C^2 convex, and that F_1 is C^1. Assume the following:

1. $f_1(x)$ is continuously differentiable in an open subset, $D \subset \Re^n$ and let $J_1(x)$ be Lipschitz continuous on D, of some rank γ_1.
2. There exists $x_* \in D$, such that $\nabla f(x_*) = 0$.

   Assume that on the identification neighborhood $N(x_*, r) \subset D$, that the following are satisfied:
3. J_1 is bounded on $N(x_*, r)$.
4. There exists $\sigma \geq 0$, such that, for all $x \in N(x_*, r)$, for any choice of generalized Jacobian, $J_2(x)$, $J_2(x_*)$,

$$\|(J_1(x) - J_1(x_*))^T F_1(x_*) + (J_2(x) - J_2(x_*))^T F_2^+(x_*)\| \leq \sigma\|x - x_*\|. \tag{6.1.3.1}$$

5. λ, the smallest eigenvalue of $J^T(x_*)J(x_*)$, for any choice of generalized Jacobian, satisfies:

$$\lambda > \sigma.$$

Then there exists $c \in (1, \frac{\lambda}{\sigma})$, and there exists $\epsilon > 0$ such that for all $x \in N(x_*, \epsilon)$ and for any choice of generalized Jacobian

$$\|(J^T(x)J(x))^{-1}\| \leq \frac{c}{\lambda},$$

and for all $x_0 \in N(x_*, \epsilon)$, the Gauss-Newton iterates

$$x_{k+1} = x_k - (J(x_k)^T J(x_k))^{-1} J(x_k)^T F(x_k)$$

are well-defined, converge to x_* and satisfy, for some $\gamma > 0$ and for all α where $\|J(x)\| \leq \alpha$, for all $x \in N(x_*, r)$:

$$\|x_{k+1} - x_*\| \leq \frac{c\sigma}{\lambda}\|x_k - x_*\| + \frac{c\alpha\gamma}{2\lambda}\|x_k - x_*\|^2, \tag{6.1.3.2}$$

$$\|x_{k+1} - x_*\| \leq \frac{c\sigma + \lambda}{2\lambda}\|x_k - x_*\| < \|x_k - x_*\|. \tag{6.1.3.3}$$

Proof. Since any generalized Jacobian satisfies that $\|(J(x_k)^T J(x_k))^{-1}\| \leq \frac{1}{\lambda}$ and since $c > 1$, then Proposition 1.3.1 assures a neighborhood of x_*, such that $\|(J^T(x)J(x))^{-1}\| \leq \frac{c}{\lambda}$. By 1, since J_1 is Lipschitz continuous on $N(x_*, r)$ of some rank γ, and since F_2 is C^2 convex, choose ϵ small enough to insure a Proposition 5.2.1 identification neighborhood. Then Proposition 5.2.2 holds for some γ_2. Thus, there is $\gamma = \gamma_1 + \gamma_2 > 0$, such that any generalized Jacobian of F at $x \in N(x_*, r)$ satisfies:

$$\|F(x_*) - F(x) - J(x)(x_* - x)\| \leq \frac{\gamma}{2}\|x_* - x\|^2.$$

The proof of the theorem follows directly the pattern of Theorem 1. ||

If the residual of $F(x_*)$ is zero, just like in the linear case, convergence is q-quadratic. This is similar to the result of Burke[1983] for his Gauss-Newton like projected steps. He also discusses the *Mangasarian*

Fromowitz Constraint Qualification (MFCQ) at a given point x. The *MFCQ* for a given point x is satisfied, if there exists a direction p, such that

$$\begin{aligned} &\langle \nabla F_2^i(x), p\rangle < 0 \quad \text{where } i \in I^0(x) \cup I^+(x), \\ &\langle \nabla F_1^i(x), p\rangle = 0 \quad \text{where } i = 1, \ldots, m1, \\ &\text{and } \{\nabla F_1^i(x) | i = 1, \ldots, m1\} \quad \text{are linearly independent.} \end{aligned} \tag{$MFCQ$}$$

Robinson showed that if a point x satisfies the MFCQ condition, then the condition is satisfied for all points in some neighborhood of x. Using this result and a result of Daniel[1973], Burke showed that his Gauss-Newton directions, starting from a point in an identification neighborhood of a stationary point x_* having $f(x_*) = 0$, where the MFCQ holds for some other point $\bar{x}$ in this neighborhood and $f(\bar{x}) = 0$, then the iterates converge q-quadratically to x_*.

The local convergence behavior of Levenberg-Marquardt directions for the general convex case is comparable to Gauss-Newton. The applicability of either of these methods depends on how nonlinear the function F is at x^* and on the size of the residual at x_*, as is discussed in section (4.1.1).

6.2 Global Convergence Results

Both global convergence results are applicable as in the case of linear inequalities. The Armijo line search follows directly without adaptation, since $\nabla f_2(x)$ is Lipschitz continuous of some rank γ on the lower level set of the first iterate, since F_2 is convex and C^2. By assuming that $\nabla f_1(x)$ is Lipschitz continuous of some rank on the lower level set of the first iterate, then the Armijo line search conditions are satisfied.

Similarly, the global trust region method using generalized Hessian model function also follows directly. Since F_2 is convex and C^2, then, in an identification neighborhood of any accumulation point, the Hessian model function converges quadratically to f_2. By assuming that F_1 is C^2, the conditions of the generalized Hessian model for the global trust region method are satisfied.

7 Applications

The least square algorithm developed in previous sections can be used to solve the assignment problem subproblem (2.5.1) as well as other common optimization problems, such as linear programming problems and the linear complementarity problem. In section 7.1, we develop algorithms to solve the piece-wise linear assignment problem in section 2. The algorithm calls subroutines based on Han's algorithm for solving systems of linear inequalities in a least square sense and the algorithms developed in section 4 for solving systems of linear inequalities and nonlinear equations. The local minimum stationary point returned as a solution to the subproblem (2.5.2) may not yield a global minimum of the least square formulation. We show that non-global local minimum stationary points are either associated with a w-component vertex which is not a possible $\{0, 1\}$ solution, or there is no solution with integral w-component. In the latter case, one can form cutting planes and reformulate the relaxed problem. Based on the assumption that proper starting points can be found for each iteration (corresponding to vertices of the w-component where $w_j = 0$ or 1), the algorithm terminates finitely in a solution with integral w-component or finds that no solution exists. The global convergence property of the least square algorithms in section 4 gives the potential for fewer than 2^n iterations. In section 7.2, we show how the the relaxed linear programming problem can be formulated as system of linear inequalities and equalities which can be solved using Han's algorithm. Combined with the nonlinear equations, the problem represents the peice-wise linear assignment problem subproblem (2.5.1) which can be solved by the algorithm developed in section 7.1.

7.1 Solving the Piece-wise Linear Assignment Problem

Recall that a solution w, y, z of the relaxed primal generated by r, u, v which solves the dual problem in (2.2) is a solution to a system of linear inequalities and equalities with respect to the following index sets :

$$J^{=} := \{j \mid r_j \in (0,1)\} \iff w_j\phi_j = \sum_i c_{i,j} y_{i,j} = z_j$$

$$J^{\geq} := \{j \mid r_j = 1\} \implies \sum_i c_{i,j} y_{i,j} = z_j \geq w_j\phi_j$$

$$J^{\leq} := \{j \mid r_j = 0\} \iff \sum_i c_{i,j} y_{i,j} \leq w_j\phi_j = z_j$$

$$I_j^0 := \{i \mid -v_{i,j} + u_i - c_{i,j} r_j < 0\} \iff y_{i,j} = 0$$

$$I_j^w := \{i \mid v_{i,j} > 0\} \iff y_{i,j} = w_j.$$

Here w, y, z satisfy:

$$\begin{aligned}
y_{i,j} &= 0, \quad \forall i \in I_j^0, \forall j \\
y_{i,j} &\geq 0, \quad \forall i, j \\
\sum_j y_{i,j} &= 1, \quad \forall i \\
y_{i,j} &= w_j, \quad \forall i \in I_j^w, \forall j \\
y_{i,j} &\leq w_j, \quad \forall i, j \\
w_j &\leq 1, \quad \forall j \\
t_j(1 - w_j) &= 0, \quad \forall j \\
\sum_i c_{i,j} y_{i,j} &= z_j = w_j\phi_j, \quad \forall j \in J^{=} \\
\sum_i c_{i,j} y_{i,j} &= z_j \geq w_j\phi_j, \quad \forall j \in J^{\geq} \\
\sum_i c_{i,j} y_{i,j} &\leq z_j = w_j\phi_j, \quad \forall j \in J^{\leq}.
\end{aligned} \tag{7.1.1}$$

Let $P := \{(w, y, z) \mid$ which satisfies (7.1.1)$\}$ be the nonempty convex polyhedral relaxed primal solution set generated by a single dual solution. (Note that z is an auxillary variable and would not be used in the computation.) Consider P as the non-empty convex polytope in $\Re^{n+mn}$ (by eliminating the variable z). All points in P have $w_j \in [0,1]$. As in section 2.5, use the single variable x to represent the variable (w, y). Define the function F_1 to represent the nonlinear forcing function of w, where $(F_1)_j : \Re \mapsto \Re$, and $(F_1(x))_j := w_j(1 - w_j)$. Let the matrix A and the vector b represent the linear transformations in the inequalities, and the matrix C and the vector d represent the linear transformations in the equalities. The least square algorithm for solving systems of non-linear equations and linear inequalities returns a local minimum stationary point x_* which solves in a least square sense:

$$\begin{aligned}
F_1(x) &= 0, \\
Ax &\leq b, \\
Cx &= d.
\end{aligned} \tag{7.1.2}$$

Now, define $F : \Re^{n+mn} \mapsto \Re^p$, where $p = m + 4n + 2mn + \sum_j(|I_j^0| + |I_j^w|)$. (The notation $|I_j^0|$ means the cardinality of the set I_j^0.) Then

$$F(x) := \begin{pmatrix} F_1(x) \\ (Ax-b)_+ \\ (Cx-d) \end{pmatrix} = \begin{pmatrix} F_1(x) \\ F_2(x) \\ F_3(x) \end{pmatrix}.$$

Let $f : \Re^{n+mn} \mapsto \Re$ be defined:

$$\begin{aligned} f(x) &:= f_1(x) + f_2(x) + f_3(x), \text{ where} \\ f_1(x) &:= \frac{1}{2}\langle F_1(x), F_1(x)\rangle, \\ f_2(x) &:= \frac{1}{2}\langle (Ax-b)_+, (Ax-b)_+\rangle, \\ f_3(x) &:= \frac{1}{2}\langle (Cx-d), (Cx-d)\rangle. \end{aligned}$$

Observe that the w-component vertices of $[0,1]^n$ are determined by $w_j = 0$ or 1 for each j. A convex subset of $[0,1]^n$ associated with each vertex is determined by $0 \le w_j < \frac{1}{2}$ or $\frac{1}{2} < w_j \le 1$, for each j. There are 2^n of these subsets, one associated with each w-component vertex. The minimum of $f_1(x)$ on each of these subsets is attained at the associated w-component vertex. Points having $w_j = \frac{1}{2}$ for some j are not identified with any of the vertices. We show the following property about local minimum stationary points of f.

Proposition 7.1.

If x_* is a local minimum stationary point of f, then x_* satisfies:

1.) If $x_* \in P$, then $w_j^* \in \{0, \frac{1}{2}, 1\}$, for all j. If $f(x_*) = 0$, then $w_j^* \in \{0, 1\}$, for all j.

2.) If $x_* \notin P$ and if $w_j^* = \frac{1}{2}$ for some j, then for all $x \in P$, the only feasible choice is $w_j = \frac{1}{2}$ for these j.

Proof. The first result uses the property that $\nabla f_2(x_*) + \nabla f_3(x_*) = 0$, for $x_* \in P$. The solutions of $\nabla f_1(x) = 0$ satisfy:

$$w_j^*(1-w_j^*)^2 - (w_j^*)^2(1-w_j^*) = 0,$$

which implies $w_j^* \in \{0, 1, \frac{1}{2}\}$. Then if $f(x_*) = 0$, it must be that $w_j^* \in \{0, 1\}$.

For the second result, in order that x_* is a local minimum stationary point, then $\nabla f(x_*) = 0$. If there exists α, β such that $0 \le \alpha \le \frac{1}{2} \le \beta \le 1$, $\alpha \ne \beta$, and $[\alpha, \beta]$ is feasible for w_j, then x_* is not a local minimum stationary point. Rather, the w_j^* component is a local maximum stationary point with respect to this component. Otherwise, suppose there exists $\alpha \le \beta$, where either $\alpha, \beta \in [0, \frac{1}{2})$ or $\alpha, \beta \in (\frac{1}{2}, 1]$ and $[\alpha, \beta]$ is feasible for w_j. Consider:

$$(\nabla f(x))_j = (w_j - \alpha) + w_j^2(1-w_j) - w_j(1-w_j)^2$$

evaluated at $w_j = \frac{1}{2}$,

$$= (\frac{1}{2} - \alpha) < 0.$$

This means x_* with $w_j^* = \frac{1}{2}$ is not a local minimum stationary point. Similarly,

$$(\nabla f(x))_j = (w_j - \beta) + w_j^2(1 - w_j) - w_j(1 - w_j)^2$$

evaluated at $w_j = \frac{1}{2}$,

$$= (\frac{1}{2} - \beta) > 0.$$

Again, this means x_* with $w_j^* = \frac{1}{2}$ is not a local minimum stationary point. Because of the convexity of P, these are the only possible types of feasible regions, and so the only possibility for x_* to be a local minimum stationary point with $w_j^* = \frac{1}{2}$ for some j is that the only feasible choice for these w_j is $w_j = \frac{1}{2}$. ||

This proposition assures that local minimum stationary points with any $w_j^* = \frac{1}{2}$ represent a condition where there is no point with w integral in P. Further, any other local minimum stationary point has all $w_j^* \neq \frac{1}{2}$. Thus, in the case that $f(x_*) \neq 0$, one can identify a vertex of the w-component associated with this point, or if any of the $w_j^* = \frac{1}{2}$, then no solution with integral w-component exists. It follows in the case where $f(x_*) \neq 0$ where no $w_j^* = \frac{1}{2}$ that the corresponding w-component vertex is not the w-component vertex of a possible solution.

The cutting plane method to solve integer programming problems originated in the work of Gormory[1960,1963] and is described in elementary texts such as Garfinkel and Nemhauser[1972]. Conn and Cornuéjols[1990], Cornuéjols et al[1989], and Ahn et al[1988] describe the cutting planes that generate facets of the uncapacitated facility location polytope. This polytope is the convex hull of the solution set of:

$$\begin{aligned} \min \quad & \sum_j z_j \\ \text{subject to:} \quad & \\ w_j - y_{i,j} &\geq 0, \quad \forall i,j \\ -\phi_j w_j + z_j &\geq 0, \quad \forall j \\ -\sum_i c_{i,j} y_{i,j} + z_j &\geq 0, \quad \forall j \\ \sum_j y_{i,j} &= 1, \quad \forall i \\ y_{i,j} &\geq 0, \quad \forall i,j \\ w_j &\in \{0,1\}, \quad \forall j. \end{aligned} \tag{7.1.3}$$

The following cutting plane:

$$y_{o,l} + y_{p,h} + y_{p,l} + y_{q,l} + y_{q,k} + y_{o,k} - w_h - w_l - w_k \leq 1 \tag{7.1.4}$$

defines a facet of the uncapacitated facility location polytope for any $h, l, k \in J$, and $o, p, q \in I$, such that $h \neq l \neq k$ and $o \neq p \neq q$. It cuts off fractional basic solutions of (2.1.2) the relaxed primal solution, where all the variables in (7.1.4) take the value $\frac{1}{2}$. The authors show examples of using these special cutting planes, where, after just a few iterations, a $\{0,1\}$ solution is feasible.

We describe an algorithm to find a $\{0,1\}$ solution for the problem in section 2. Define $LLS(P)$ to be Han's algorithm which returns a least square solution x^* of P define by (7.1.1), as well as residuals r_*. Define $NLLS(f)$ to be the algorithm which returns a local minimum stationary point, x_*, of f generated by (7.1.2), which is a least square solution to a system of non-linear equations and linear inequalities. The algorithm calls both LLS(P), NLLS(f). Consider $l \in V$, where initially $V = \{1, 2, \ldots, 2^n\}$ represents the vertices of the w-component of x. For each $l \in V$, choose an initial starting point with the associated w-component vertex, selecting $w_j = 0$ or 1 appropriately. If P contains points having w-component in this subset of $[0,1]^n$ associated with the particular w-component vertex, then if one chooses a starting point in P, identified with this w-component vertex, then the algorithm NLLS(f) should converge to a point with the same identified w-component vertex. Thus, one way to choose the starting points for NLLS(f) is to call Han's algorithm to find a feasible point in P and associating $w_j = 0$ or 1 depending on the closest endpoint. However, after several iterations, one may not find different vertices based on points in P. Another approach is to choose a starting point based on a particular w-component vertex where the y-component of the point x_0 satisfies the conditions in (7.1.1) in a least square sense. Then, call NLLS(f) with the starting point and test the conditions of Proposition 7.1. If $f(x_*) = 0$, then the algorithm has found a solution where the w component is integral. If any of the $w_j^* = \frac{1}{2}$, then the algorithm stops and using cutting planes described by Conn and Cornuéjols[1990], a new problem is formulated, feasibility is restored, and a new set P is formed. Then, one can again run the algorithm below with this new set P. Lastly, if $x_* \notin P$, then none of the $w_j^* = \frac{1}{2}$ and there is an associated w-component vertex of x_*. If this vertex is the same as that of the initial point, then this vertex is removed from the list V. If the the initial vertex is different, then, since the initial starting point was chosen with its w-component in this subset of $[0,1]^n$ with y-component as close as possible to feasibility, then one can remove this initial starting point vertex from the list as well. Because of the convexity of P, if there are any points of P having w-components in this subset, then the algorithm NLSS should find this local minimum with its w-component within this subset. Finally, the algorithm iterates selecting another vertex from the list V. If at any iteration, both endpoints of some w_j are found to be infeasible, then Gomory cutting planes would be developed to cut off non-integral solutions, the relaxed program reformulated and using a new set P, the algorithm would be rerun.

The algorithm is shown below:

Algorithm using NLLS and LLS :

$V = \{1, 2, \ldots, 2^n\}$;

P is the set defined by (7.1.1) or by cutting planes;

ITER:

SELECT l from V; (Choose vertex from $V \neq \emptyset$)

SET x_0; (calling $LLS(P_y)$ with set w-component and y-component as variable.)

$x_* = NLLS(f)$; (where f is the function defined by the set P.)

IF $f(x_*) = 0$, THEN done; (x_* is an integral solution)

IF ANY $w_j^* = \frac{1}{2}$, THEN quit; (Use cutting planes to form new relaxed program and new set P)

(Thus, x_* is infeasible and no $w_j^* = \frac{1}{2}$)

l = w-component vertex of x_0;

$V = V\backslash\{l\}$; (Remove initial vertex from V.)

$l_* =$ w-component vertex of x_*;

IF $l \neq l_*$, THEN $V = V\backslash\{l_*\}$;

FOR j = 1 TO n; BEGIN;

  IF $w_j = 0$ or 1 both infeasible THEN quit;

  (Use Gomory cutting planes to form new relaxed program and new set P.)

  END;

GO TO ITER;

This algorithm terminates finitely if an integral w-component solution exists in the set P. At each iteration, at least one different vertex is removed from the list. The above algorithm chooses all starting points for calling NLSS based on an associated w-component vertex, setting the $w_j = 0$ or 1 depending on the vertex. If a point in P has this w-component vertex, then LSS algorithm would find a solution with this integral w-component. If not, then the point returned from LSS gives the least square solution in the y-component having this associated w-component vertex. If there are points in P having w-component in the subset associated with this w-component vertex, then NLSS (using this starting point) finds the local minimum stationary point with the same associated w-component vertex. If the starting point is the least square solution of the y-component with a given w-component vertex, then if the NLSS algorithm converges to a point having a different w-component vertex, then one can eliminate both w-component vertices from consideration. Thus, it is possible that two different vertices are removed at an iteration. If after sufficient number of possible w-component vertices are eliminated, one finds for some j such that $w_j = 0$ or 1 are both eliminated, or if the NLSS algorithm finds a local minimum stationary point with some $w_j^* = \frac{1}{2}$, then no point in the set P has integral w-component. One would then form cutting planes and a new set P and rerun the algorithm.

This algorithm calls LSS at each iteration in finding starting points and calls NLSS once each iteration. The global convergence property of the algorithms in section 4 provides the potential that fewer than 2^n iterations are needed.

7.2 Solving Linear Programming Problems

Han[1980], Mangasarian[1981], and Stewart[1987] describe using iterative continuous algorithms to solve linear programs. Specifically, Han's algorithm could be used to solve a linear complementarity problem which takes . e same form as the solution set of a linear program. One might also use this algorithm to solve the relaxed linear program described in section 2. Using the duality property, the optimal solution of the primal problem (2.1.2) and the optimal solution of the dual problem (2.2.1) must satisfy:

$$\sum_j z_j \geq \sum_i u_i - \sum_j t_j,$$

in addition to satisfying all the constraints in both problems. Therefore, one could formulate a convex set, P, in the space of the combined primal and dual variables, where P satisfies:

$$\begin{aligned}
\sum_j z_j &\geq \sum_i u_i - \sum_j t_j, \\
w_j - y_{i,j} &\geq 0, \quad \forall i, j \\
-\phi_j w_j + z_j &\geq 0, \quad \forall j \\
-\sum_i c_{i,j} y_{i,j} + z_j &\geq 0, \quad \forall j \\
\sum_j y_{i,j} &= 1, \quad \forall i \\
y_{i,j} &\geq 0, \quad \forall i, j \\
w_j &\leq 1, \quad \forall j \\
\sum_i v_{i,j} - \phi_j s_j &= 0, \quad \forall j \\
-v_{i,j} + u_i - c_{i,j} r_j &\leq 0, \quad \forall i, j \\
s_j + r_j &= 1, \quad \forall j \\
v_{i,j} &\geq 0, \quad \forall i, j \\
s_j &\geq 0, \quad \forall j \\
r_j &\geq 0, \quad \forall j \\
t_j &\geq 0. \quad \forall j
\end{aligned} \tag{7.2.1}$$

The solution set P represents all the solutions to the relaxed linear programming problem, unlike using a single dual variable to generate a reduced solution set. One may consider using P generated by (7.2.1) and the algorithm to solve for an integral solution. This avoids the iterations needed to find other dual solutions, since the entire solution set of the relaxed problem is characterized by (7.2.1).

8 Examples of Algorithm

For the example problem, the strategy as suggested in 7.2 is used to characterize the full set of solutions of the dual problem and search within this set for a $\{0, 1\}$ solution. Recall from section 2, the full set of solutions of the example problem satisfy: $y_{i,j} = 0$, except as follows:

$$y_{1,1} = w_1 = 1, \quad y_{2,3} = y_{4,3} = w_3 = 1, \quad y_{3,3} = \frac{3}{44},$$

and $w_2, y_{3,2}$, and $y_{3,1}$ satisfy:

$$w_2 \le 1, \quad y_{3,1} \ge \frac{6}{40},$$
$$y_{3,2} = \frac{41}{44} - y_{3,1}, \quad 0 \le y_{3,2} \le w_2 \le \frac{4}{3} y_{3,2}.$$

One can simplify, eliminating the variable $y_{3,1}$ which is fully determired, and expre^s as a problem in two variables, w and y:

$$\begin{aligned} -y &\le 0, \\ y - \frac{43}{55} &\le 0, \\ y - w &\le 0, \\ w - \frac{4}{3} y &\le 0, \\ w - 1 &\le 0. \end{aligned} \tag{8.1}$$

Expressing $f(x) = f(w, y)$:

$$\begin{aligned} f(x) = \frac{1}{2}(w^2(1-w)^2 &+ (-y)_+^2 + (y - \frac{43}{55})_+^2 \\ &+ (y-w)_+^2 + (w - \frac{4}{3}y)_+^2 + (w-1)_+^2). \end{aligned} \tag{8.2}$$

See the contour plot of $f(x)$ shown in figure 3 below, where one observes there are local minina stationary points at $w = 0, y = 0$, and $w = 1, y \in [\frac{3}{4}, \frac{43}{55}]$.

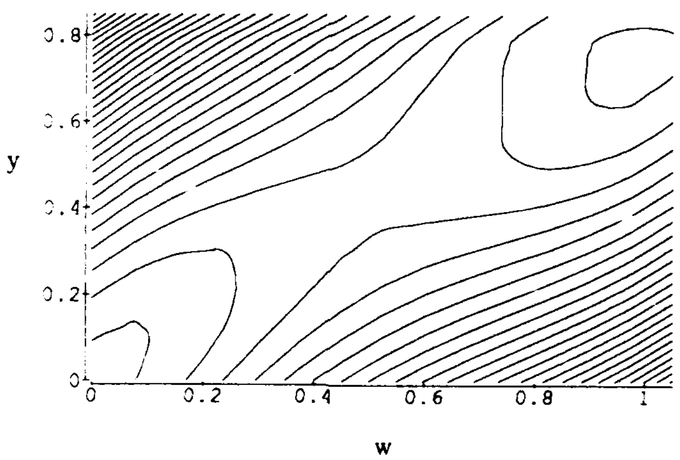

Figure 3. Contour Plot of f(x)

For this problem,

$$A = \begin{pmatrix} 0 & -1 \\ 0 & 1 \\ -1 & 1 \\ 1 & -\frac{4}{3} \\ 1 & 0 \end{pmatrix}, \qquad b = \begin{pmatrix} 0 \\ \frac{43}{55} \\ 0 \\ 0 \\ 1. \end{pmatrix}$$

An example using the Gauss-Newton like algorithm to solve (7.3.2) is illustrated, using initial starting point $x_0 = (w, y)^T = (1, 1)^T$.

Initial Iteration: $f(x_0) = (\frac{12}{55})^2$, $I^0(x_0) = \{3, 5\}$, $I^-(x_0) = \{1, 4\}$, and $I^+(x_0) = \{2\}$.

$$\nabla f(x_0) = \begin{pmatrix} 0 \\ \frac{12}{55} \end{pmatrix}, \quad J_1(x_0) = \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix}, \quad J_2(x_0) = \begin{pmatrix} 0 & 0 \\ 0 & 1 \\ -\alpha & \alpha \\ 0 & 0 \\ \beta & 0 \end{pmatrix},$$

where $\alpha \in [0, 1]$ and $\beta \in [0, 1]$. Selecting the generalized Jacobian $J_2(x_0) \in \partial(Ax_0 - b)_+$, where $\alpha = \beta = 0$, one calculates:

$$J^T(x_0)J(x_0) = J_1^T(x_0)J_1(x_0) + J_2^T(x_0)J_2(x_0) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad d_0 = -\nabla f(x_0) = \begin{pmatrix} 0 \\ -\frac{12}{55} \end{pmatrix}.$$

Iteration 1: Choosing stepsize $\lambda_0 = 1$, then

$$x_1 = \begin{pmatrix} 1 \\ \frac{43}{55} \end{pmatrix}, \quad \nabla f(x_1) = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

so x_1 solves (7.3.2), with $f(x_1) = 0$. Therefore, x_1 also solves the assignment problem.

If a different generalized Jacobian, $J_2(x_0) \in \partial(Ax_0 - b)_+$, is chosen where $\alpha = \beta = 1$ above, then

$$J^T(x_0)J(x_0) = \begin{pmatrix} 3 & -1 \\ -1 & 2 \end{pmatrix} \quad \text{and} \quad d_0 = -(J^T(x_0)J(x_0))^{-1}\nabla f(x_0) = \begin{pmatrix} -.0436 \\ -.1309 \end{pmatrix}.$$

Iteration 1a: Choosing stepsize $\lambda_0 = 1$ gives

$$x_1 = \begin{pmatrix} 0.9564 \\ 0.8691 \end{pmatrix}, \quad \nabla f(x_1) = \begin{pmatrix} -0.0381 \\ 0.0873 \end{pmatrix}.$$

$I^0(x_1) = \emptyset$, $I^-(x_1) = \{1, 3, 4, 5\}$, and $I^+(x_1) = \{2\}$. Notice that for $x_* = (1, \frac{43}{55})^T$, that x_1 is in an identification neighborhood of x_*, where $I^0(x_*) = \{2, 5\}$, $I^-(x_*) = \{1, 3, 4\}$, and $I^+(x_*) = \emptyset$. Since $I^0(x_1) = \emptyset$, then $J_2(x_1) \in \partial(Ax_1 - b)_+$ is uniquely determined:

$$J^T(x_1)J(x_1) = \begin{pmatrix} 0.8331 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad d_1 = -\begin{pmatrix} 0.0457 \\ -0.0873 \end{pmatrix}.$$

Choosing stepsize $\lambda_1 = 1$ gives Iteration 2a:

$$x_2 = \begin{pmatrix} 1.0021 \\ 0.7818 \end{pmatrix}, \quad \nabla f(x_2) = \begin{pmatrix} 0.0042 \\ 0 \end{pmatrix}.$$

$I^0(x_2) = \{2\}$, $I^-(x_*) = \{1, 3, 4\}$, and $I^+(x_*) = \{5\}$, Choosing $J_2(x_2) \in (Ax_2 - b)_+$, then

$$J_2(x_2) = \begin{pmatrix} 0 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \\ 1 & 0 \end{pmatrix} \quad \text{gives} \quad J^T(x_2)J(x_2) = \begin{pmatrix} 2.0084 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad d_2 = \begin{pmatrix} -\frac{1}{6} \\ 0 \end{pmatrix}.$$

Choosing stepsize $\lambda_2 = 1$ so that

$$d_2 = \begin{pmatrix} -0.0021 \\ 0 \end{pmatrix} \quad \text{gives} \quad x_3 = \begin{pmatrix} 1 \\ \frac{43}{55} \end{pmatrix}, \quad f(x_3) = 0, \quad \nabla f(x_3) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

9 Conclusions

We have developed methods based on continuous optimization to solve a special type of assignment problem, having piece-wise linear additive separable cost functions. This problem is not easily solved using combinatoric methods, due to the nonlinear cost function. Further, common continuous optimization methods do not handle the $\{0, 1\}$ contraints. The strong relaxation of the $\{0, 1\}$ constraints yields a linear program or, equivalently, a system of inequalities and equalities. Adding nonlinear equations to this system to force a $\{0, 1\}$ solution yields a problem which we showed can be solved using least square methods. We developed algorithms extending methods such as Gauss-Newton, Levenberg-Marquardt, Newton's, and quasi-Newton, for deriving search directions, and global convergence algorithms based on the model trust region method and inexact line search. We have developed the details of these new algorithms and an overall theory for local and global convergence for both linear inequalities as well as systems of inequalities of convex C^2 functions. This work in non-differentiable optimization is based on on fundamental properties of generalized differentiability. We showed that these properties are useful in developing theoretical convergence results of these new algorithms to solve a special type of problem, least square solution of finite systems of inequalities. The local convergence theory uses the special approximating property that the Taylor series based on the generalized differential constructs possesses in an identification neighborhood of a stationary point. We developed and verified new approximating results for both the case of linear inequalities and systems of inequalities of convex C^2 functions, as well as the details of selecting generalized differential constructs. From this, we developed algorithms based on Gauss-Newton, Levenberg-Marquardt and Newton's methods, using generalized differential constructs and showed that local convergence properties are similar to the smooth case. We developed a comparable method based on quasi-Newton updates, defined a bounded deterioration property for generalized Hessians, and, assuming this property, verified local q-linear convergence.

The global convergence of the Armijo line search, which we showed can be used directly with no modification, depends on the Lipschitz property of the gradient of the convex least square function on lower level sets. We extended the global trust region method to handle systems of inequalities by using a model function based on generalized Hessians and verified convergence.

We developed an algorithm to solve the original assignment problem using the algorithm for solving systems of nonlinear equations and linear inequalities. This overall algorithm either finds a $\{0, 1\}$ solution within the relaxed problem solution set, or it determines that no such solution exists, thus requiring cutting planes be added to the relaxed problem formulation. We developed and verified a property of local minimum stationary points of the least square formulation which, along with proper choice of starting points at each iteration, assures that the assignment problem algorithm converges, that is, either a solution is found, or it is confirmed that none exists. We showed how to formulate the relaxed problem as a system of linear inequalities and equalities which can be solved using Han's algorithm.

This work demonstrates strong solvability properties by using the Euclidean norm for regularizing a non-differentiable function into a differentiable function, as well as the increased solvability of transforming a discontinuous problem into a continuous problem. As shown, the least square formulation is numerically stable to small perturbations of computational inaccuracies. The least square formulation for solving the system of inequalities and equalities yields a solution even if the system is infeasible. The least square formulation, although not equivalent to the original assignment problem, yields important information about local minimum stationary points of the least square objective function for the assignment problem. Further, the least square formulation for solving systems of inequalities provides a useful tool for other applications, such as solving linear programs and linear and non-linear complementarity problems.

Further work would address more general functions, especially non-convex functions. Also, further work should treat the invertibility assumptions made for local convergence results, deriving the specific conditions, such as the MFCQ, which assure invertibility, similar to Burke[1983], Harker and Pang[1990], and Harker and Xiao[1990]. In addition, the specific conditions which assure the bounded deterioration property for the secant update in the quasi-Newton algorithm need to be developed.

Lastly, the theoretical algorithms developed herein should be implemented and tested on large scale problems, giving a basis for practical comparison with other algorithms and comparison with smooth problems.

Bibliography

S. Ahn, C. Cooper, G. Cornuéjols, A. Frieze, "Probabilistic Analysis of a Relaxation for the k-Median Problem", *Mathematics of Operations Research* **13** (1988), 1-31.

R. Bramley, "QR Column Pivoting Versus SVD", Unpublished Manuscript, Center for Supercomputing Research and Development, University of Illinois, Urbana-Champaign, Spring 1992.

J. V. Burke, "Methods for Solving Generalized Systems of Inequalities with Application fo Nonlinear Programming", Ph.D. Dissertation, University of Illinois, Urbana-Champaign, 1983.

J. V. Burke, and J. J. Moré, "On the Identification of Active Constraints", *SIAM Journal on Numerical Analysis* **25** (1988), 1197-1211.

J. V. Burke, J. J. Moré, and G. Toraldo, " Convergence Properties of Trust Region Methods for Linear and Convex Constraints", *Mathematical Programming* **47** (1990), 305-336.

F. H. Clarke, *Optimization and Nonsmooth Analysis* , Les Publications CRM, Université de Montréal, 1989.

A. R. Conn and G. Cornuéjols, "A Projection Method for the Uncapacitated Facility Location Problem," *Mathematical Programming* **46** (1990), 272-298.

A. R. Conn, N. I. M. Gould, and Ph. L. Toint, "Global Convergence of a Class of Trust Region Algorithms for Optimization with Simple Bounds", *SIAM Journal on Numerical Analysis* **25** (1988), 433-460.

G. Cornuéjols, G. L. Nemhauser, and L. A. Wosley, "The Uncapacitated Facility Location Problem", in R. L. Frances and P. Mirchandani, eds., *Discrete Location Theory* , Wiley, New York, 1989.

J. E. Dennis, R. B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, Prentice-Hall, New Jersey, 1983.

J. E. Dennis and H. F. Walker, "Convergence Theorems for Least Change Secant Update Methods, *SIAM Journal on Numerical Analysis* **18** (1981), 949-987.

R. Fletcher, *Practical Methods of Optimization*, Wiley, New York, 1987.

P. E. Gill, W. Murray, M. Wright, *Practical Optimization* , Academic Press, New York, 1981.

R. S. Garfinkel, G. L. Nemhauser, *Integer Programming*, Wiley, New York, 1972.

R. E. Gomory, "Solving Linear Programming Problems in Integers," *Proceedings of the Symposium on Applied Mathematics* **10** (1960), 211-215.

R. E. Gomory, "An Algorithm for Integer Solutions to Linear Programs", in R. L. Graves and P. Wolfe, eds., *Recent Advances in Mathematical Programming*, McGraw-Hill, New York, 1963, 269-302.

S. P. Han, "Least Square Solution of Linear Inequalities", Unpublished Manuscript, University of Wisconsin at Madison, 1982.

P. T. Harker, J. -S. Pang, "Finite-dimensional Variational Inequality and Nonlinear Complementarity Problems: A Survey of Theory, Algorithms, and Applications," *Mathematical Programming* **48** (1990), 161-220.

P. Harker, B. Xiao, "Newton's Method for the Nonlinear Complementarity Problem: A B-differentiable Equation Approach," *Mathematical Programming* **48** (1990), 339-357.

J. -B. Hiriart-Urruty, "A New Set-Valued Second Order Derivative for Convex Functions", *FERMAT Days 85: Mathematics for Optimization*, Elsevier, North Holland, 1986.

C. Lemarechal, "Constructing Bundle Methods for Convex Optimization", *FERMAT Days 85: Mathematics for Optimization*, Elsevier, North Holland, 1986.

O. L. Mangasarian, "Iterative Solution of Linear Programs," *SIAM Journal on Numerical Analysis* **18** (1981), 606-614.

J. M. Ortega, W. C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Equations*, Academic Press, New York, 1970.

M. Padberg, B. Rinaldi, "A Branch-and-Cut Algorithm for the Resolution of Large Scale Symmetric Travelling Salesman Problems," *SIAM Review* **35** (1991), 60-100.

S. M. Robinson, "Stability Theory of Systems of Inequalities. Part I: Linear Systems," *SIAM Journal on Numerical Analysis* **12** (1975), 754-769.

S. M. Robinson, "Stability Theory of Systems of Inequalities. Part II: Differentiable Nonlinear Systems," *SIAM Journal on Numerical Analysis* **13** (1976), 497-513.

S. M. Robinson, "Newton's Method for a Class of Nonsmooth Functions," *SIAM Journal on Numerical Analysis* **13** (1990) .

R. T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, New Jersey, 1979.

H. Scramm, J. Zowe, "A Version of the Bundle Idea for Minimizing a Nonsmooth Function: Conceptual Idea, Convergence Analysis, Numerical Results", *SIAM Journal of Optimization* **2** (1992), 121-152.

G. W. Stewart, "An Iterative Method for Solving Linear Inequalities", TR-1833, Dept. of Computer Science and Institute for Physical Science and Technology, Univ. of Maryland, College Park MD, 1987.

J. M. Varah, "Relative Sizes of the Hessian Terms in Nonlinear Parameter Estimation", *SIAM Journal on Scientific and Statistical Computing* **11** (1990), 174-179.

DISTRIBUTION

Chief of Engineers
ATTN: CEHEC-IM-LH (2)
ATTN: CEHEC-IM-LP (2)
ATTN: CERD-L

Fort Belvoir, VA 22060
ATTN: CECC-R

Defense Technical Info. Center 22304
ATTN: DTIC-FAB (2)

8
11/92